
**Präpositionen und Deutsch als Fremdsprache: Quantitative
Fallstudien im Lernerkorpus MERLIN**

Tassja Weber

Tassja Weber

Präpositionen und Deutsch als Fremdsprache: Quantitative Fallstudien im Lernerkorpus
MERLIN

**Zugleich Dissertation der Universität Mannheim
(Tag der Disputation: 21.11.2019)**

Dekan: Prof. Dr. Philipp Gassert
Erste Gutachterin: Prof. Dr. Angelika Storrer
Zweiter Gutachter: Prof. Dr. Stefan Engelberg



Danksagung

Ich möchte mich bei all den Menschen bedanken, die dazu beigetragen haben, dass ich meine Promotion erfolgreich abschließen konnte und meine Promotionszeit in guter Erinnerung behalten werde.

Mein größter Dank gilt meiner Betreuerin Prof. Dr. Angelika Storrer: Durch sie habe ich die Chance erhalten, an der Universität Mannheim arbeiten und promovieren zu dürfen. Ich bedanke mich für ihre fortwährende Unterstützung in dieser Zeit, ihr Vertrauen in mich und meine Arbeit, ihre konstruktive Kritik, und ihre Beratung – sei es in individuellen Einzelgesprächen oder im Rahmen von regelmäßigen Kolloquien, in denen ich meine Arbeit präsentieren konnte.

Einen großen Dank möchte ich ebenfalls meinem Zweitgutachter und Mitbetreuer Prof. Dr. Stefan Engelberg aussprechen – in seinem Kolloquium habe ich meine Arbeit und Fortschritte präsentieren und diskutieren dürfen. Dieser Austausch hat mich stets zur Reflexion meiner eigenen Arbeit angeregt und sie auf diese Weise positiv beeinflusst.

Des Weiteren bedanke ich mich bei Sascha Wolfer, der mich bei den statistischen Auswertungen beraten hat. Ich danke ihm sehr für die Zeit, die er sich dafür genommen hat. Ebenfalls möchte ich den Mitarbeiterinnen und Mitarbeitern des Instituts für Angewandte Sprachforschung von EURAC Research in Bozen (Italien) meinen Dank aussprechen – allen voran Andrea Abel, Aivars Glaznieks, Alexander König, Egon Stemle und Lionel Nicolas. Im Rahmen meines Forschungsaufenthalts bei EURAC Research haben sie mir in verschiedener Form dabei geholfen, den Fokus meiner Dissertation weiter zu schärfen und methodische Herausforderungen zu bewältigen.

Einen herzlichen Dank auch an alle Projektbeteiligten von MERLIN – jenem digitalen, frei verfügbaren Lernerkorpus, ohne das meine Dissertation nicht in dieser Form möglich gewesen wäre. An dieser Stelle möchte ich erneut das Institut für Angewandte Sprachforschung von EURAC Research erwähnen, das diese Datensammlung derzeit pflegt und zugänglich macht. Vielen Dank!

Meinen Kolleginnen am Lehrstuhl für Germanistische Linguistik der Universität Mannheim während meiner Promotionszeit gebührt ebenfalls ein herzliches Dankeschön – für die Unterstützung, das Feedback zu meinen Vorträgen und den vielfältigen Austausch in- und außerhalb von Mittagspausen und Forschungskolloquien. Ich schätze diese angenehme Atmosphäre am Lehrstuhl wirklich sehr! Ein großes Dankeschön auch an die ehemaligen studentischen Hilfskräfte am Lehrstuhl – allen voran Janna Berauer und Christina Grimm, die mich bei den Annotationen der Forschungsdaten unterstützt haben.

Im privaten Umfeld gilt mein größter Dank meinem langjährigen Partner Sebastian und unseren Familien: Ihre bedingungslose Liebe und Unterstützung hat mir in meiner Promotionszeit viel Kraft und Energie gegeben.

Mannheim, März 2020

Inhaltsverzeichnis

Einleitung	1
Teil I: Grundlagen	5
1. Präpositionen und Präpositionalphrasen	5
1.1 Präpositionen	6
1.1.1 Kasusrektion	7
1.1.2 Sprachliche Funktion und semantischer Gehalt.....	9
1.1.3 Verschmelzung der Präposition mit definitivem Artikel	15
1.1.4 Zusammenfassung der zentralen Terminologie I	16
1.2 Präpositionalphrasen (PP)	17
1.2.1 PP unter syntaktischer Perspektive: Syntaktische Funktionen	18
1.2.2 Zusammenfassung der zentralen Terminologie II	32
1.2.3 PP unter valenzgrammatischer Perspektive	33
1.2.4 Zusammenfassung der zentralen Terminologie III	41
1.2.5 Andere Verwendungsweisen von PP.....	42
1.2.6 Zusammenfassung der zentralen Terminologie IV.....	45
1.3 Präpositionen und Präpositionalphrasen aus kontrastiver Perspektive	45
1.4 Zusammenfassung.....	48
2. Aktuelle theoretische Ansätze zum Zweit- und Fremdspracherwerb.....	50
2.1 Grundlegende Konzepte gebrauchsbasierter Ansätze zum L2-Erwerb.....	51
2.1.1 Form-Funktionsverbindungen.....	51
2.1.2 Assoziatives Lernen	52
2.1.3 Rationale Sprachverarbeitung.....	53
2.1.4 Exemplar-basiertes Lernen	53
2.1.5 Emergente Relationen und Muster	54
2.1.6 Zwischenfazit I	54
2.2 Merkmale assoziativen Lernens im Kontext des L2-Erwerbs	55
2.2.1 Overshadowing und Attention Blocking.....	56
2.2.2 Zwischenfazit II.....	58
2.3 Methodologie gebrauchsbasierter Ansätze	58
2.4 Zusammenfassung.....	59
3. Lernerkorpusforschung.....	60
3.1 Einordnung der Lernerkorpusforschung	61
3.2 Das Lernerkorpus: Begriffsbestimmung.....	63
3.2.1 Korpusaufbau und Metadaten.....	64
3.2.2 Typologie	66
3.3 Annotationen in einem Lernerkorpus.....	69
3.3.1 Automatische Annotationen: Wortarten.....	70
3.3.2 Manuelle Annotationen: Zielhypothese und Fehlerannotation.....	72
3.4 Zentrale Analysemethoden der Lernerkorpusforschung	79
3.4.1 Computerunterstützte Fehleranalyse	79
3.4.2 Kontrastive Interlanguage-Analyse	84
3.5 Schnittstellen der Lernerkorpusforschung.....	89
3.5.1 Korpuslinguistik.....	89
3.5.2 Maschinelle Sprachverarbeitung.....	90
3.5.3 Zweit-/Fremdsprachenunterricht.....	91

3.5.4 Zweit-/Fremdspracherwerbsforschung.....	93
3.6 Zusammenfassung.....	99
4. Skizzierung des Forschungsstands.....	100
4.1 Präpositionen und PP in der deutschen Zweit- und Fremdspracherwerbsforschung.....	100
4.1.1 Überblick.....	101
4.1.2 Erstsprache der Lernenden.....	103
4.1.3 Sprachkompetenz	109
4.1.4 Syntaktische Funktion der PP.....	112
4.1.5 Semantischer Gehalt der Präposition.....	115
4.1.6 Zusammenfassung.....	117
4.2 Präpositionen und PP in der (deutschen) Lernerkorpusforschung.....	117
4.2.1 Einleitung.....	117
4.2.2 Untersuchungen zum Gebrauch von Präpositionen und PP in deutschen Lernerkorpora	118
4.2.3 Arbeiten zur automatischen Erkennung fehlerhafter Präpositionen	120
4.2.4 Zusammenfassung.....	122
5. Zusammenfassung Teil I.....	123
Teil II: Empirische Untersuchungen	128
6. Datengrundlage: Das Lernerkorpus MERLIN	128
6.1 Korpus-Aufbau und Daten	128
6.2 Annotationen	135
6.2.1 Zentrale Annotationsebenen: Überblick.....	135
6.2.2 Zielhypothesen.....	137
6.2.3 Fehlerannotationen.....	138
6.3 Dateiformat und Korpus-Zugang	141
6.4 Zusammenfassung und Motivation der Korpus-Wahl für die Dissertation	141
7. Fallstudien im Lernerkorpus MERLIN: Überblick.....	142
7.1 Analysemethoden	143
7.2 Hinweise zur statistischen Auswertung	146
7.2.1 Regressionsanalyse	146
7.2.2 Spezifische Anmerkungen zur Regressionsanalyse in der vorliegenden Dissertation	150
8. Fallstudie 1: Erstsprache der Lernenden	153
8.1 Einleitung	153
8.2 Untersuchungsgegenstand und Leitfragen	153
8.3 Datenauswahl und Datenerhebung.....	156
8.4 Datenaufbereitung	160
8.5 Datenauswertung.....	160
8.5.1 Analysemethoden	161
8.5.2 Ergebnisse.....	162
8.6 Zusammenfassung und Diskussion der Gesamtergebnisse.....	181
9. Fallstudie 2: Syntaktische Funktion der PP	184
9.1 Einleitung	184
9.2 Untersuchungsgegenstand und Leitfragen	185
9.3 Datenauswahl und Datenerhebung.....	187

9.4 Datenaufbereitung	191
9.4.1 Annotationskategorien.....	191
9.4.2 Annotationsprozess.....	194
9.5 Datenauswertung	195
9.5.1 Analysemethoden	199
9.5.2 Ergebnisse.....	199
9.6 Zusammenfassung und Diskussion der Gesamtergebnisse	213
10. Fallstudie 3: Semantischer Gehalt der Präposition	217
10.1 Einleitung	217
10.2 Untersuchungsgegenstand und Leitfragen	218
10.3 Datenauswahl und Datenerhebung	220
10.4 Datenaufbereitung	223
10.4.1 Annotationskategorien.....	223
10.4.2 Annotationsprozess.....	227
10.5 Datenauswertung	229
10.5.1 Analysemethoden	231
10.5.2 Ergebnisse.....	232
10.6 Zusammenfassung und Diskussion der Gesamtergebnisse	252
11. Zusammenfassung Teil II und Ausblick	256
Abkürzungsverzeichnis	263
Abbildungsverzeichnis	263
Tabellenverzeichnis	263
Anhang 1: Outputs der statistischen Auswertungen	287
Anhang 2: Annotationshandbuch	291

Einleitung

Eine (Fremd-)Sprache zu lernen, heißt u.a. auch, sprachliche Merkmale der Zielsprache (L2) im Input wahrzunehmen und sie zu verarbeiten, sprachliche Muster und Regularitäten aus dem Input abzuleiten, und (zielsprachlich) korrekt anzuwenden. Zu den grammatischen Merkmalen der Zielsprache Deutsch, die aus zweit- und fremdspracherwerbstheoretischer Perspektive besonders interessant sind, zählen u.a. Präpositionen (Grießhaber 2010: 293). Präpositionen (z. B. *an, auf, bei, mit, für*) gehören zu den häufigsten Wörtern des Deutschen (Duden 2016: 612, Grießhaber 2007a: 636) und repräsentieren damit einen zentralen Lerngegenstand im Erwerb des Deutschen als L2. Präpositionen verbinden sich vorrangig mit nominalen Einheiten zu Phrasen (z. B. *an der Wand, für jemanden*), deren Kasus sie regieren und die beispielsweise abhängig vom syntaktischem Kontext Restriktionen im Gebrauch der entsprechenden Präposition aufweisen (vgl. *warten auf/*für, liegen auf/hinter/bei*) und die hinsichtlich ihrer Bedeutung mehr oder weniger transparent sind (vgl. *in den Wald, in drei Wochen, sich in jemanden verlieben*) (Duden 2016: 615). Eine Präposition zu produzieren bedeutet folglich, eine Präpositionalphrase zu produzieren. Der Lerngegenstand *Präposition* umfasst somit auch immer den Lerngegenstand *Präpositionalphrase (PP)*.

In der Forschung gelten Präpositionen und PP als „eine große Schwierigkeit“ (Fandrych/Thurmair 2018: 189) und „komplexer Lerngegenstand“ (Turgay 2010a: 1) für Deutschlernende. Turgay (2010a: 1) macht deutlich, dass „für die Bildung einer Präpositionalphrase [...] die Verwendung einer Präposition und, wenn diese auftritt, die lexikalische Kenntnis dieser Präposition nötig [sind] und darüber hinaus die Rektion der Präposition“ und damit einhergehend die Realisierung entsprechender Kasusmarkierungen. Nach Hoffmann (2016: 612) gehören Präpositionen zu „mögliche[n] Lernprobleme[n] im Bereich der deutschen Grammatik“, und das unabhängig von der Erstsprache der Lernenden. Die Schwierigkeit des Lerngegenstands sieht Hoffmann (2016: 612) vor allem in der Bedeutung der Präposition, ihrer Kasusrektion und ihrer Bindung an spezifische Verben. Zu den „eigentlichen“ Problemen im Erwerb und Gebrauch der L2 Deutsch zählt nach Auffassung von Grießhaber (2010: 306) jedoch die Realisierung einer zielsprachlich korrekten Präposition, also der Gebrauch der Präposition innerhalb einer PP. Hier zeigen sich viele Unsicherheiten, wie die folgenden authentischen Beispiele¹ aus schriftlichen Texten von Lernenden des Deutschen als Fremdsprache (DaF) illustrieren:

1. *Ich meine, wenn man nach Ausland fährt, dann muss man Informationen suchen über die neue Kultur.*
2. *Ich werde Ihnen dankbar sein, wenn Sie meine Fragen antworten könnten.*
3. *Ich kann nicht bei dir besuchen, weil ich Deutsch lernen muss.*

Spezifische Unsicherheiten im Gebrauch von Präpositionen und PP manifestieren sich bei DaF-Lernenden nicht nur in der Wahl der zielsprachlichen korrekten Präposition (1), sondern auch

¹ Die Beispiele stammen aus dem Lernerkorpus MERLIN, das ebenfalls als Datengrundlage der vorliegenden Dissertation dient. Die entsprechenden Identifikationsnummern (ID) lauten: 1. ID: 1031_0003219, 2. ID: 1023_0109891, 3. ID: 1061_0120274.

in Auslassungen von zielsprachlich geforderten Präpositionen bzw. PP (2) und in Hinzufügungen von Präpositionen bzw. PP in Kontexten, in denen zielsprachlich keine Präposition bzw. PP gefordert ist (3). Bisherige Forschungen zum Erwerb und Gebrauch von Präpositionen und PP durch DaF-Lernende liefern wertvolle Erkenntnisse zu möglichen Einflussvariablen im Gebrauch von Präpositionen und PP, fokussieren sich jedoch auf qualitative Auswertungen und nehmen i.d.R. einzelne Lernende in den Fokus. Digitale Lernerkorpora – d. h. digital aufbereitete, mit linguistischen Annotationen angereicherte und elektronisch zugängliche Sammlungen mit (schriftlichen) Äußerungen von Fremdsprachenlernenden (vgl. Granger 2008) – werden hierbei erst seit kurzem und bisher nur vereinzelt eingesetzt. Untersuchungen zum Gebrauch von Präpositionen bzw. PP unter Rückgriff auf Daten aus Lernerkorpora bei gleichzeitiger Berücksichtigung aktueller spracherwerbstheoretischer Annahmen und unter Anwendung statistischer Analyseverfahren stellen ein Desiderat dar, dem sich die vorliegende Arbeit annehmen möchte.

Die vorliegende Dissertation widmet sich dem Gebrauch von Präpositionen und PP durch DaF-Lernende und untersucht korpusgestützt, welche ausgewählten Variablen den Gebrauch von Präpositionen und PP durch DaF-Lernende beeinflussen. Der Schwerpunkt liegt dabei auf den Unsicherheiten im Gebrauch von Präpositionen und PP. Als Datengrundlage dient das digitale, frei verfügbare und mit vielfältigen Annotationen angereicherte Lernerkorpus MERLIN (Abel et al. 2014). Das Lernerkorpus enthält schriftliche Produktionen von DaF-Lernenden unterschiedlicher Kompetenzniveaus und eignet sich damit besonders gut für die quantitative Untersuchung und Skizzierung kompetenzbedingter Entwicklungen im Sprachgebrauch. Unter Berücksichtigung grundlegender Annahmen aktueller gebrauchsbasierter Ansätze zur Relevanz von Form-Funktionsverbindungen im L2-Erwerb werden Hypothesen zu ausgewählten Einflussvariablen im Gebrauch von Präpositionen und PP aufgestellt und in quantitativ ausgerichteten Fallstudien im Lernerkorpus MERLIN überprüft. Die ausgewählten Einflussvariablen sind: Die Erstsprache der Lernenden, die Sprachkompetenz, die syntaktische Funktion der PP und der semantische Gehalt der Präposition. Im Fokus steht lediglich der Gebrauch der Präpositionen (innerhalb von PP), die Realisierung der Kasusreaktion wird nicht untersucht.

Ziel der Arbeit ist es, empirisch fundierte und statistisch gesicherte Erkenntnisse zum Einfluss ausgewählter Variablen auf den Gebrauch von Präpositionen und PP durch DaF-Lernende unter Berücksichtigung spracherwerbstheoretischer Annahmen zu liefern und auf diese Weise das Potenzial von Lernerkorpora für die quantitative (DaF-) Spracherwerbsforschung zu illustrieren. Somit möchte ich in der vorliegenden Arbeit exemplarisch aufzeigen, wie Annahmen aus der Spracherwerbsforschung anhand von Lernerkorpusdaten überprüft und mögliche Einflussvariablen im Sprachgebrauch von Lernenden ermittelt werden können. Die Arbeit bettet sich in die Lernerkorpusforschung ein und ist innerhalb dieser an der Schnittstelle zur Zweit- und Fremdspracherwerbsforschung zu verorten.

Die Dissertation gliedert sich in zwei Teile: Im ersten Teil behandle ich die terminologischen Grundlagen sowie die Verortung der Dissertation im Forschungsfeld. Im

zweiten Teil gehe ich auf die quantitativen Fallstudien im Lernerkorpus MERLIN ein, wobei jeder ausgewählten Einflussvariable eine eigene Fallstudie gewidmet ist.

Der erste Teil umfasst vier Großkapitel, in denen die terminologischen Grundlagen sowie der Forschungsstand erarbeitet werden. Zunächst wird das untersuchte Phänomen – Präpositionen und PP im Deutschen – vorgestellt und die Terminologie für die Dissertation herausgearbeitet (Kap. 1). Hierbei orientiere ich mich größtenteils an der aktuellen Auflage der Duden-Grammatik (Duden 2016, 9. Auflage). Im Vordergrund des Kapitels stehen unterschiedliche Perspektiven auf Präpositionen und PP. Fokussiert werden die syntaktische, die valenzgrammatische sowie die sprachkontrastive Perspektive auf Präpositionen und PP. In diesem Zusammenhang werden die zentralen Charakteristika ausgewählter Präpositionen und PP thematisiert, die für die empirischen Fallstudien im Lernerkorpus zentral sind. Im Einzelnen sind dies die Präpositionsreaktion und der semantische Gehalt von Präpositionen in ausgewählten syntaktischen Funktionen von PP (vor allem Objekt-PP und adverbiale PP), die Verbindung zwischen Form und Funktion bei Präpositionen und PP sowie Äquivalente zu deutschen Präpositionen und PP in anderen typologisch verschiedenen bzw. typologisch ähnlichen Sprachen. Im Anschluss (Kap. 2) gehe ich auf grundlegende Konstrukte und Annahmen aktueller theoretischer gebrauchsbasierter Ansätze zum Zweit- und Fremdspracherwerb ein, die die empirischen Untersuchungen im Lernerkorpus MERLIN leiten. Zentral werden hier die Mechanismen des assoziativen Lernens skizziert und der Einfluss von Form-Funktionsverbindungen der Erstsprache sowie der Einfluss von Form-Funktionsverbindungen der Zielsprache näher erläutert. Zum Abschluss gehe ich auf die Methodologie gebrauchsbasierter Ansätze ein und motiviere den Einsatz von Lernerkorpora als geeignete Datengrundlage im Kontext gebrauchsbasierter Ansätze zum Zweit- und Fremdspracherwerb. Im 3. Kapitel gebe ich einen Überblick über die Lernerkorpusforschung, in die sich die vorliegende Dissertation einbettet. In diesem Kapitel wird u.a. der Begriff *Lernerkorpus* näher bestimmt. Der Fokus des Kapitels liegt auf spezifischen (manuellen) Annotationen in einem Lernerkorpus – der Zielhypothese und der Fehlerannotation – sowie auf den zentralen Analysemethoden der Lernerkorpusforschung – der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse – die in der vorliegenden Dissertation Anwendung finden. Ebenfalls gehe ich näher auf die Schnittstellen der Lernerkorpusforschung zu anderen Forschungsfeldern ein, wobei die Schnittstelle zur Zweit- und Fremdspracherwerbsforschung als zentrale Schnittstelle für die vorliegende Dissertation im Vordergrund steht. Im letzten Kapitel des ersten Teils (Kap. 4) wird der Forschungsstand zum untersuchten Phänomen skizziert. Hier werden in Unterkapiteln Forschungserkenntnisse aus der deutschen Zweit- und Fremdspracherwerbsforschung sowie der (deutschen) Lernerkorpusforschung präsentiert. Im Fokus des Kapitels stehen insgesamt Erkenntnisse zu ausgewählten Einflussvariablen im Gebrauch deutscher Präpositionen und PP sowie die Herausarbeitung methodologischer und inhaltlicher Forschungslücken. Ebenfalls zeige ich anhand von Entwicklungen in der (englischen) Lernerkorpusforschung an der Schnittstelle zur maschinellen Sprachverarbeitung, dass linguistische Untersuchungen in einem Lernerkorpus wertvolle Anknüpfungspunkte für Ansätze zur Entwicklung automatischer Fehlererkennungssysteme für die Zielsprache Deutsch bieten. Zum Abschluss des ersten Teils

sowie als Überleitung zum zweiten Teil der Dissertation fasse ich in Kapitel 5 die zentralen Punkte des ersten Teils zusammen.

Der zweite Teil der Dissertation präsentiert die quantitativen Fallstudien im Lernerkorpus MERLIN und bildet damit das Herzstück der vorliegenden Arbeit. Die ausgewählten Einflussvariablen im Gebrauch von Präpositionen und PP durch DaF-Lernende werden jeweils in separaten Fallstudien behandelt. Allen Fallstudien liegt das Lernerkorpus MERLIN (Abel et al. 2014) zugrunde, das im Kapitel 6 im Detail vorgestellt und typologisiert wird. MERLIN ist ein trilinguales Lernerkorpus und enthält u.a. ein Subkorpus mit schriftlichen Produktionen von DaF-Lernenden. Dieses Subkorpus dient als Datengrundlage der vorliegenden Dissertation. Im Anschluss an die Beschreibung der Datengrundlage gebe ich einen Überblick über die Fallstudien und die jeweils untersuchten Einflussvariablen (Kap. 7). Analysemethoden, die in allen Fallstudien Anwendung finden, werden in Kapitel 7.1 genauer beschrieben. In Kapitel 7.2 gehe ich näher auf die statistische Auswertung mittels (logistischer) Regressionsanalysen ein und erläutere die einzelnen Parameter der Regressionsanalysen sowie die entsprechenden Darstellungskonventionen in der vorliegenden Arbeit. Die empirischen Untersuchungen zu ausgewählten Einflussvariablen werden in Kapitel 8, 9 und 10 in einzelnen Fallstudien behandelt. Die Fallstudien sind nach einem ähnlichen Muster aufgebaut: Zunächst stelle ich den Untersuchungsgegenstand vor und präsentiere im Anschluss die zentralen Leitfragen und Hypothesen, die jeweils empirisch im Lernerkorpus MERLIN überprüft werden. Daran anschließend präsentiere ich die jeweilige Datengrundlage bzw. die Datenauswahl sowie die Datenerhebung und gehe näher auf die Datenaufbereitung ein. In der anschließenden Datenauswertung werden die Analysemethoden erläutert und die Ergebnisse der statistischen Auswertungen im Lichte der eingangs formulierten Leitfragen und Hypothesen interpretiert. Zum Schluss erfolgt jeweils eine Zusammenfassung und Diskussion der zentralen Ergebnisse der jeweiligen Fallstudie. In der Zusammenfassung des zweiten Teils (Kap. 11) werden die Ergebnisse aller Fallstudien in einem Überblick zusammengefasst und im Hinblick auf die übergeordneten Fragestellungen diskutiert. Zuletzt gibt Kapitel 11 eine Zusammenfassung der Arbeit sowie einen Ausblick.

Die vorliegende Dissertation fokussiert lediglich einen kleinen, ausgewählten Teilbereich des Gebrauchs von Präpositionen und PP durch DaF-Lernende. Dieser Teilbereich – der Gebrauch der Präposition ohne Berücksichtigung der Kasusreaktion – wird jedoch, wie oben bereits erwähnt, als Hauptproblem im Gebrauch von PP durch Deutschlernende aufgefasst. Die Arbeit intendiert nicht, verallgemeinernde Aussagen über den Erwerb von Präpositionen und PP zu machen, sondern konzentriert sich auf ausgewählte Variablen, deren Einfluss auf den Gebrauch von Präposition und PP durch Deutschlernende unterschiedlicher Kompetenzniveaus empirisch untersucht wird. Die Erkenntnisse der einzelnen Fallstudien können lediglich in Bezug auf die analysierten Daten interpretiert werden sind daher nicht zu verallgemeinern. Ebenfalls verfolgt die vorliegende Arbeit nicht das Ziel, zur Grammatikschreibung des Deutschen beizutragen.

Die Fallstudien in dieser Dissertation und deren empirisch fundierte und statistisch gestützte Erkenntnisse sollen erste Anknüpfungspunkte für die weitere, erwerbstheoretisch informierte Erforschung von Einflussvariablen im DaF-Spracherwerb und -Sprachgebrauch

bieten und das Potenzial von Lernerkorpora als Ressource für die Spracherwerbsforschung deutlich machen.

Abschließend möchte ich an dieser Stelle noch auf die Publikationsform der vorliegenden Dissertation eingehen. In der deutschen Lernerkorpusforschung wird der freien Verfügbarkeit von wissenschaftlichen Ergebnissen und Forschungsdaten ein hoher Stellenwert beigemessen. Die meisten mir bekannten deutschen Lernerkorpora sind (z.T. nach Anfrage) frei verfügbar. Hierzu gehören die Lernerkorpora MERLIN (Abel et al. 2014), Falko (Lüdeling et al. 2008), CLEG (Maden-Weinberger 2015), KANDEL (Vyatkina 2016), ALeSKo (Zinsmeister/Breckle 2012), das Kobalt-Korpus (Zinsmeister et al. 2012) sowie die sich im Aufbau befindenden Lernerkorpora Dulko (Hirschmann/Nolda 2019) und DiSKo (Wisniewski 2018a, Universität Leipzig 2017), die nach ihrer Fertigstellung frei verfügbar sein werden. Die (kosten)freie Verfügbarkeit von MERLIN hat die vorliegende Arbeit und damit die empirischen Fallstudien zu Einflussvariablen im Gebrauch von Präpositionen und PP durch Deutschlernende erst möglich gemacht. Durch die freie Verfügbarkeit von MERLIN ist gleichzeitig auch die Nachvollziehbarkeit und Reproduzierbarkeit der empirischen Ergebnisse gewährleistet. Diesem Leitgedanken folgend habe ich mich dazu entschlossen, meine Dissertation² über die Universitätsbibliothek Mannheim in einem digitalen Open-Access-Format zu publizieren und sie damit kostenfrei und dauerhaft zugänglich zu machen.

Teil I: Grundlagen

1. Präpositionen und Präpositionalphrasen

Das vorliegende Kapitel beschreibt die für meine Untersuchungen relevanten Eigenschaften von Präpositionen und Präpositionalphrasen (PP) in der deutschen Sprache. Behandelt werden zunächst die zentralen Charakteristika von Präpositionen (Kap. 1.1). Daran anschließend wird die Präpositionalphrase behandelt (Kap. 1.2). Hierbei stehen vor allem deren Struktur und Funktion im Vordergrund.

Der Gegenstand Präpositionen und PP wird aus unterschiedlicher sprachtheoretischer Perspektive in vielen Publikationen behandelt (s. z. B. Philippi/Tewes 2008, Hoffmann 2016, GDS 1997, Jacobs 2003: 384f., Rostila 2014, 2018, Proost 2015, Wiese/Pohle 2016, Marossek 2016). Ziel des vorliegenden Kapitels ist es nicht, diese unterschiedlichen sprachtheoretischen Perspektiven aufzuarbeiten bzw. sie zu kontrastieren (vgl. dazu z. B. Elsen 2014). Mit dem vorliegenden Kapitel verfolge ich das Ziel, eine terminologisch einheitliche Grundlage für meine empirischen Untersuchungen in einem Lernerkorpus zu schaffen. Bei der Beschreibung des Untersuchungsgegenstands orientiere ich mich primär an der aktuellen Duden-Grammatik (Duden 2016, 9. Auflage). Im Kontext der DaF-Forschung nimmt die Duden-Grammatik eine wichtige Rolle ein, denn sie intendiert „auf die gezielten Bedürfnisse von Lehrenden und Lernenden des Deutschen als Fremdsprache einzugehen, sodass sich die Grammatik gewinnbringend im Unterricht und im Selbststudium einsetzen lässt“ (Duden

² Hierzu zählt ebenfalls das Annotationshandbuch sowie die annotierten Forschungsdaten.

2016: Vorwort). Die Wahl der Duden-Grammatik als primäre Grundlage des vorliegenden Kapitels orientiert sich somit am Schwerpunkt der Dissertation, in der der Sprachgebrauch von DaF-Lernenden aus einem gesteuerten Erwerbskontext (Sprachunterricht) ausgewertet und der Einfluss ausgewählter Variablen auf den Gebrauch von Präpositionen und PP untersucht wird.

1.1 Präpositionen

Präpositionen wie *auf, in, unter, mit, von, dank, kraft, mithilfe* sind nicht flektierbar und zählen zu den am häufigsten vorkommenden Wortarten des Deutschen³. Hinsichtlich des genauen Inventars liegen abhängig vom Zählverfahren und der Wortartendefinition schwankende Zahlen vor (vgl. Duden 2016: 612, Eisenberg 2013: 187, Grießhaber 2007a: 639f.). Bei Präpositionen handelt es sich um eine offene Wortklasse, sodass sich der Bestand dieser Klasse ständig erweitert (vgl. Duden 2016: 612/GDS 1997: 2075⁴). Unter den Präpositionen gibt es z.T. erhebliche Frequenzunterschiede hinsichtlich ihres Vorkommens in der deutschen Sprache, wobei z. B. die Präpositionen *in, an* und *für* deutlich häufiger im Sprachgebrauch vorkommen als z. B. die Präpositionen *kraft* und *angesichts* (vgl. Duden 2016: 612).

Weiterhin unterscheiden sich Präpositionen hinsichtlich ihrer internen Struktur bzw. strukturellen Komplexität. So differenziert der Duden (2016: 612) zwischen primären Präpositionen (1.), sekundären Präpositionen (2.) sowie tertiären Präpositionen (3.). Somit können Präpositionen sowohl aus einfachen Wortformen als auch aus mehreren Wortformen bestehen:

1. *in, auf, zu, um*
2. *zufolge, mithilfe, aufgrund, dank, kraft*
3. *in Bezug auf, in Anbetracht*

Primäre Präpositionen (1.) stellen die prototypischen Vertreter der Wortartenkategorie dar. Solche Präpositionen gehören zur ältesten Schicht und damit zum „Kernbestand“ der Präpositionen (Grießhaber 2007a: 631, Eisenberg 2013: 185). Sekundäre Präpositionen (2.) sind historisch betrachtet jünger als primäre Präpositionen. Diese Präpositionen entwickelten sich aus Substantiven bzw. substantivischen Gruppen (vgl. Grießhaber 2007a: 631, Duden 2016: 613), was z. T. noch heutzutage deutlich zu erkennen ist (vgl. *dank, kraft* bzw. *mithilfe, aufgrund*)⁵. Als tertiäre Präpositionen (3.) werden diejenigen Präpositionen gezählt, die aus mehreren Bestandteilen zusammengesetzt sind. Einer dieser Bestandteile ist wiederum eine (primäre) Präposition.

Prototypische Präpositionen stehen, wie die Etymologie der Bezeichnung nahelegt, vor dem sprachlichen Element, auf das sie sich beziehen (Bezugselement) (4.). Im Deutschen gibt es zudem Postpositionen, d. h. nachgestellte Präpositionen, (5.) sowie Zirkumpositionen, also

³ Laut Duden (2016: 612) findet sich in „fast jedem Satz“ eine Präposition.

⁴ Grießhaber (2007a: 631) spricht im Gegensatz dazu von einer geschlossenen Wortklasse.

⁵ Bei einigen Präpositionen dieser Art ist der Entstehungsprozess hin zur einer Präposition noch nicht vollständig abgeschlossen, Zusammen- und Getrennschreibung existieren parallel (vgl. *auf Grund/aufgrund, an Stelle/anstelle*) (Duden 2016: 613).

‚umschließende‘ bzw. umrahmende Präpositionen (6.). Sowohl Post- als auch Zirkumpositionen sind im Vergleich zu den prototypischen vorangestellten Präpositionen sehr selten (Duden 2016: 612)⁶.

4. auf dem Tisch

5. dem Autor zufolge

6. um Gottes willen

Im Fokus dieser Arbeit stehen primäre Präpositionen als prototypische Vertreter ihrer Klasse. Sekundäre und tertiäre Präpositionen werden in der vorliegenden Arbeit nicht betrachtet. Aus diesem Grund wird an dieser Stelle auf eine ausführliche Darstellung der sekundären und tertiären Präpositionen verzichtet. Mit dem Begriff *Präposition* wird in der vorliegenden Arbeit ausschließlich auf vorangestellte Präpositionen verwiesen, da diese im Fokus dieser Arbeit stehen. Im Folgenden wird näher auf die typischen Charakteristiken von Präpositionen im Deutschen eingegangen.

1.1.1 Kasusreaktion

Unter *Rektion* versteht man allgemein die Festlegung grammatischer Merkmale einer Einheit durch ein Wort (Duden 2016: 779). Handelt es sich bei diesem grammatischen Merkmal um den Kasus, spricht man auch von *Kasusreaktion* (Duden 2016: 779). Unter dem Begriff *Kasusreaktion* verstehe ich hier das Festlegen, d. h. Regieren, des Kasus durch eine Präposition. Rektion verweist damit auf eine Art der Abhängigkeit. Präpositionen legen den Kasus ihres Bezugselements fest, damit liegt eine Abhängigkeit vor zwischen Präposition und Kasus des Bezugselements. Die Kasusreaktion ist auf Akkusativ, Dativ und Genitiv beschränkt, eine Präposition regiert nie den Nominativ (vgl. Duden 2016: 6178, Eisenberg 2013: 182). Der Dativ gilt dabei als „Normalkasus“ (Duden 2016: 849) bei Präpositionen.

Der Großteil der Präpositionen regiert genau *einen* festen Kasus (7.). Einige Präpositionen sind in der Lage, zwei Kasus zu regieren (8.)⁷. Zu den Präpositionen, die zwei Kasus regieren können, gehören hauptsächlich die sogenannten Wechselprepositionen (9.), die abhängig von der Syntax und Bedeutung der Prepositionalphrase den Dativ oder den Akkusativ regieren können (Weiteres s. Exkurs unten). Aber auch Präpositionen, die den Dativ und Genitiv regieren können, zählen zu dieser Gruppe (10.). Während die Unterscheidung zwischen Dativ vs. Akkusativ bei Wechselprepositionen syntaktisch-semantic bestimmt wird und die Unterschiede im Kasus mit funktionalen Unterschieden einhergehen, ist bei den entsprechenden Präpositionen die Wahl des Dativs und Genitivs eher stilistisch bedingt (vgl. Eisenberg 2013: 182, Duden 2016: 619).

7. *mit, bei, für, aus, angesichts, anstelle, etc.*

⁶ Die GDS (1997: 2085) nimmt z. B. an, dass im Deutschen lediglich eine „einzige eindeutige“ Zirkumposition existiert, und zwar *um [...] willen*.

⁷Es existieren ebenfalls Präpositionen ohne eindeutige Kasusreaktion (z. B. *entlang*) (vgl. GDS 1997: 44, Eisenberg (2013): 182, Duden 2016: 615, Gießhaber 2007a: 148).

8. *auf, unter, in, über, etc.*
9. *in, an, auf, vor, über, unter, etc.*
10. *wegen, statt, laut, etc.*

Bei Präpositionen, bei denen die Kasusreaktion eher stilistisch bedingt ist, geht die Verwendung der Kasus mit keiner Bedeutungsveränderung der Aussage einher, hier spricht man auch von *Nebenkasus* (Duden 2016: 618, Helbig/Buscha 2001a: 180). Die Ausdifferenzierung zweier (semantisch äquivalenter) Kasus ist hier primär im Sprachwandel begründet (Duden 2016: 619). Bei anderen Präpositionen kann man von einem *Doppelkasus* sprechen, da die Präpositionen zwei Kasus (Dativ und Akkusativ) regieren, und diese auch mit funktionalen Unterschieden korrelieren (Duden 2016: 620). Die folgende Tabelle (Tab. 1) gibt eine zusammenfassende Übersicht zur Kasusreaktion von Präpositionen.

	Präpositionen mit einem (festen) Kasus			Präpositionen mit zwei Kasus	
Kasus	Dativ	Akkusativ	Genitiv	Nebenkasus (Genitiv/Dativ)	Doppelkasus (Dativ/Akkusativ)
Beispiele	<i>aus, bei, mit, seit, von, zu, nach, zu, außer</i>	<i>bis, durch, für, gegen, ohne, um</i>	<i>anhand, anstatt, aufgrund, infolge, hinsichtlich, zugunsten</i>	<i>wegen, statt, laut</i>	<i>auf, unter, in, an, neben, hinter, vor, über, zwischen</i>

Tabelle 1: Übersicht zur Kasusreaktion von Präpositionen.

Die deutsche Sprache verfügt über insgesamt neun Präpositionen mit Doppelkasus, die sowohl den Dativ als auch den Akkusativ regieren können. Man spricht in diesem Zusammenhang von Wechselpräpositionen (Duden 2016: 620, Helbig/Buscha 2001a: 182). Die neun Wechselpräpositionen sind: *auf, unter, in, an, neben, hinter, vor, über, zwischen*.

Die Kasusreaktion dieser Wechselpräpositionen wird durch die Bedeutung der Präpositionalphrase im syntaktischen Kontext gesteuert. In der Regel regieren diese Präpositionen in lokaler, statischer Verwendung den Dativ, in direktonaler Verwendung dahingegen den Akkusativ (GDS 1997: 2105, Duden 2016: 620)⁸, vgl. die folgenden Beispiele:

Die Vase steht auf dem Tisch.

Ich stelle die Vase auf den Tisch.

⁸ Der Vollständigkeit halber sei an dieser Stelle erwähnt, dass die Reaktion von Präpositionen durch Verben auch einen direkten Einfluss auf die Kasusreaktion von Wechselpräpositionen ausübt: Wechselpräpositionen, die prototypisch je nach syntaktisch-semantischer Umgebung zwei Kasus regieren, können auch durch ein Verb regiert, d. h. an dieses gebunden sein. Als regierte Präpositionen regieren Wechselpräpositionen jedoch nur *einen* (festen) Kasus, wobei die Kasusmöglichkeiten durch die Präposition vorgegeben werden (vgl. Eisenberg 2013: 79). So regiert z. B. *auf* sowohl den Dativ als auch den Akkusativ, als regierte Präposition (z. B. bei *warten auf*) aber lediglich einen der beiden Kasus, die Präposition regiert also z. B. nie den Genitiv.

Man spricht hier auch von der kasusbedingten Differenzierung zwischen (statischer) Lage und Bewegung (bzw. Richtung) und verweist diesbezüglich auf die Frageprobe mit *Wo?* vs. *Wohin?* (vgl. Duden 2016: 620), die zur Bestimmung der Kasuswahl nach Wechselpräpositionen gerne herangezogen wird. In Publikationen zur Thematik der Wechselpräpositionen, vor allem aus der Perspektive der DaF-Sprachdidaktik, ist vielerorts auf die Problematik hingewiesen worden, die die Fragen zur Ermittlung des geforderten Kasus *Wo?* und *Wohin?* bei Lernern erzeugen (vgl. Baten 2008, Wilmots/Moonen 1997, Turgay 2010b)⁹. Einen alternativen Zugang zur Erklärung der Kasusforderung nach Wechselpräpositionen liefert Leys (1989). Zur Bestimmung der Kasuswechsel bei Wechselpräpositionen verweist Leys (1989) auf ein Prinzip, das, im Gegensatz zur *Wo/Wohin*-Unterscheidung, in seiner Natur nicht-lokalistisch ist: Der Dativ zeigt laut Leys (1989: 99) ein *bestehendes* Verhältnis an, während der Akkusativ ein *entstehendes* Verhältnis kodiert. Für die Kasusreaktion der Präposition ist nach Leys (1989) ausschlaggebend, ob ein Verhältnis bereits bei der sprachlich ausgedrückten Handlung als *bestehend* gilt (dann wird der Dativ regiert) oder ob ein Verhältnis bei der Handlung erst erzeugt wird d. h. als *entstehend* verstanden wird (dann wird der Akkusativ regiert) (vgl. Leys 1989: 100). Leys (1989: 99f.) nimmt weiterhin an, dass sich die Kasusreaktion nicht nur bei Wechselpräpositionen, sondern bei allen lokalen Präpositionen auf diese Weise erklären ließe. Im Falle der Wechselpräpositionen hat man es dann mit solchen Präpositionen zu tun, die „sowohl die Vorstellung eines bestehenden als auch die Vorstellung eines entstehenden Verhältnisses zulassen“, während Präpositionen mit fester Kasusreaktion „mit nur einer von beiden Vorstellungen zu verbinden sind“ (Leys 1989: 100). Im DaF-Bereich wurde Leys' Ansatz zur Erklärung der Kasuswahl nach Wechselpräpositionen u. a. von Sylla (1999) positiv hervorgehoben. An dieser Stelle wird nicht weiter auf Wechselpräpositionen eingegangen, da die Kasusreaktion von Präpositionen nicht im Fokus der empirischen Untersuchungen liegt.

1.1.2 Sprachliche Funktion und semantischer Gehalt

Die sprachliche Funktion von Präpositionen lässt sich z. T. bereits anhand ihrer Definition in Grammatiken ableiten. So werden Präpositionen terminologisch als „Fügewörter“ (Helbig/Buscha 2001a: 177), „operative Ausdrücke“ (GDS 1997: 44), „relationierende Verfahren“ (Grießhaber 2011: 142), oder als „Verhältnis[wörter]“ (Duden 2016: 612) bzw. als „Wörter, die sprachliche Einheiten zueinander in ein Verhältnis setzen“ (Grießhaber 2007a: 629) umschrieben. In diesen Versuchen, die sprachliche Funktion von Präpositionen zu beschreiben, wird deutlich, dass die prototypische sprachliche Funktion einer Präposition darin besteht, Relationen bzw. Verhältnisse unterschiedlichster Art auszudrücken. Im Beispiel *Karl rastet auf dem Bismarckturm* stellt die Präposition *auf* eine (lokale) Relation zwischen dem Sachverhalt *Karl rastet* und *dem Bismarckturm* dar (Eisenberg 2013: 183). Präpositionen in Verwendung wie *Karl rastet auf dem Bismarckturm* wird „eine hohe kommunikative Relevanz“ zugeschrieben, „wenn es darum geht, individuelle Anschauungsräume sprachlich nachvollziehbar zu gestalten“ (Lütke 2011: 110f.). Das Beispiel illustriert die Verwendung einer Präposition als lexikalische Einheit mit eigenständig relationalem Charakter, d. h. mit einer

⁹ Siehe z. B. Balci/Kanatlı (2001) für eine gegenteilige Meinung.

eigenen Bedeutung (Eisenberg 2013: 183), die Relation ist als lokal zu klassifizieren (Duden 2016: 615). In bestimmten Verwendungskontexten verliert die Präposition dahingegen diesen „eigenständig relationalen“ Charakter (Eisenberg 2013: 183). Im Beispiel *Karl wartet auf Godot* drückt die Präposition *auf* eine komplexe Relation zwischen dem Subjekt *Karl* und dem Objekt *Godot* aus, die Präposition *auf* ist fest an das Verb gebunden und somit Bestandteil des komplexen Prädikats *warten auf* (Eisenberg 2013: 184). Die durch die Präposition ausgedrückte Relation (und damit einhergehend die Bedeutung der Präposition) wird im Duden (2016: 618) als „bedeutungsneutral“ bezeichnet.

Im Folgenden wird auf die unterschiedlichen Relationen, die Präpositionen ausdrücken können, eingegangen. Dabei folge ich Duden (2016: 615ff.) und gebrauche die Begriffe *Relation* (im Duden 2016: 815 „Verhältnis“) und *Bedeutung* synonym, d. h. wenn ich von der Relation einer Präposition spreche (z. B. lokal), meine ich damit die entsprechende Bedeutung der Präposition (z. B. lokale Präposition). Dabei ist es nicht mein Anliegen, die Bedeutung von Präpositionen im Einzelnen zu behandeln bzw. diese herauszuarbeiten. Die „teilweise sehr subtilen Unterscheidungen bei der Verwendung unterschiedlicher Präpositionen auf das gleiche Bezugs[wort] treffend und kompakt herauszuarbeiten“, stellt nach Grießhaber (2007a: 643) eine Herausforderung dar.¹⁰ Im Vordergrund dieses Kapitels steht die Skizzierung von grundlegenden Relationen, die Präpositionen ausdrücken können.

Duden (2016: 615) unterscheidet hier zunächst zwischen lokalen, temporalen, modalen und kausalen Präpositionen. Die Mehrheit der deutschen Präpositionen kann dabei je nach Gebrauch der Präpositionalphrase unterschiedliche Relationen und damit eine unterschiedliche Bedeutung ausdrücken. Im Folgenden soll kurz auf die unterschiedlichen Relationen eingegangen werden, in denen laut Duden (2016: 615ff.) die Präpositionen semantisch beschreibbar sind (lokal, temporal, modal, kausal). Diese Relationen werden abgegrenzt zur „bedeutungsneutral[en]“ Relation von Präpositionen (Duden 2016: 618). Ich möchte an dieser Stelle jedoch anmerken, dass die Relation bzw. Bedeutung einer Präposition erst im konkreten syntaktischen Kontext der entsprechenden Präpositionalphrase ermittelt werden kann, da die Klassifikation der semantisch beschreibbaren Relationen auf der Verwendung von Präpositionen innerhalb von Präpositionalphrasen in bestimmten syntaktischen Kontexten basiert. Dementsprechend ist in den folgenden Abschnitten stets ein entsprechender syntaktischer Kontext der Präpositionen bzw. Präpositionalphrase angegeben.

1.1.2.1 Lokale Relation

Die meisten primären Präpositionen sind aus Lokaladverbien entstanden (Duden 2016: 613), die lokale Bedeutung einer Präposition wird daher vielerorts als die ursprüngliche Bedeutung angenommen¹¹. Präpositionen, die lokale Relationen ausdrücken können, stellen somit den

¹⁰ Auf mögliche Ansätze zur semantischen Differenzierung von Präpositionen und ihrer Kombinierbarkeit mit Bezugswörtern geht Grießhaber (200: 642ff.) exemplarisch ein.

¹¹ Diese Auffassung leitet z.T. die Charakterisierung semantischer Rollen bei Präpositionalphrasen (vgl. Eisenberg 2013: 78f.).

„Kernbestand“ deutscher Präpositionen dar (vgl. z. B. Eisenberg 2013: 184, s. auch Duden 2016: 613).

Zu den Präpositionen, die eine lokale Relation ausdrücken können, gehören z. B. *an, auf, in, nach, hinter, um, vor, zu, gegenüber*, etc., man spricht hier auch von lokalen Präpositionen (vgl. Duden 2016: 616). Dabei können lokale Präpositionen z. B. auf eine Lage oder eine Richtung verweisen (Duden 2016: 616). Bei der Klasse der Wechselpräpositionen zeigt maßgeblich der Kasus an, ob eine Lage- oder Richtungsrelation vorliegt. Die Mehrheit der Präpositionen drückt die lokale Relation aber unabhängig von der Kasuswahl aus: So verweist *nach* z. B. auf eine Richtung (*Sie wandert nach Bozen*) und *bei* auf eine Lage (*Sie steht beim Auto*). Solchen Präpositionen wird eine inhärente Relation zugeschrieben (vgl. Duden 2016: 616). Beispiele für den Gebrauch lokaler Präpositionen sind im Folgenden aufgeführt:

Das Buch liegt auf dem Tisch
Sie nahm das Buch aus dem Regal
Sie wandert im Wald.

Lokale Präpositionen können weiterhin hinsichtlich der räumlichen Dimension, die sie ausdrücken, weiter untergliedert werden. So können lokale Präpositionen Nähe, Parallelität, eine Gegenseite oder den Bezug zu einem bestimmten Punkt spezifizieren (vgl. Duden 2016: 616).

1.1.2.2 Temporale Relation

Eine temporale Relation kann in bestimmten Kontexten z. B. durch die Präpositionen *in, für, gegen, mit, nach, über, um, vor, während* etc. ausgedrückt werden, wobei sich temporale Präpositionen auf Vor-, Nach- oder Gleichzeitigkeit des Geschehens von Ereignissen sowie Handlungen beziehen können. Sie dienen somit dazu, Geschehen hinsichtlich ihrer zeitlichen Einordnung und Ausdehnung zu spezifizieren (vgl. Duden 2016: 617), damit setzen sie Sachverhalte in zeitliche Relation zueinander (vgl. Eisenberg 2013: 186). Beispiele für den Gebrauch temporaler Präpositionen sind:

Sie kommt in drei Tagen
Sie fährt gegen Abend los
Er wird vor Einbruch der Dunkelheit eintreffen.

1.1.2.3 Kausale Relation

Kausale Relationen können in bestimmten Kontexten durch die Präpositionen *auf, aus, wegen, trotz, um, unter, mangels*, etc. hergestellt werden (vgl. Duden 2016: 617). Duden (2016: 618f.) stuft diese Präpositionen als „im weiteren Sinn kausale Präpositionen“ ein. Sie können die Relation des Grundes, des Anlasses, der Einräumung, der Einschränkung und des Zwecks kennzeichnen (Duden 2016: 617). Die folgenden Beispiele illustrieren den Gebrauch kausaler Präpositionen in einem syntaktischen Kontext:

Wegen des Regens konnte man nicht in den Garten gehen.

*Das Kind konnte vor Aufregung kaum sprechen.
Unter diesen Bedingungen mache ich nicht mit.
Sie wird für ihre gute Arbeit gelobt.*

1.1.2.4 Modale Relation

Die modale Relation wird in Duden (2016: 617) nicht näher spezifiziert und nimmt eine Art Restklasse in den Relationen ein, die Präpositionen ausdrücken können. In diese Klasse fallen Relationen, die weder als lokal, temporal oder kausal klassifiziert werden können und ebenfalls nicht zur Anzeige einer neutralen Relation verwendet werden (Duden 2016: 617). Modale Präpositionen umfassen daher „teilweise heterogene Einheiten“. Beispiele für den Gebrauch modaler Präpositionen sind im Folgenden aufgeführt:

*Der Anzug ist aus Seide.
Sie läuft diese Strecke ohne Ermüdung.
Er kauft anstelle der Schallplatte eine CD.
Otto handelt der Abmache zuwider.*

1.1.2.5 Neutrale Relation

In den oberen Abschnitten wurden Relation dargestellt, die semantisch beschreibbar sind und sich dadurch in unterschiedliche semantische Klassen einteilen lassen. Diese unterschiedlichen Klassen (lokal, temporal, kausal, modal) fassen dabei sowohl konkrete Relationen sowie davon abgeleitete Relationen zusammen. Nach Eisenberg (2013: 186) sind beispielweise „die meisten Präpositionen mit temporaler Bedeutung [...] offensichtliche Übertragungen aus dem Räumlichen.“ Die Übertragung einer temporalen Relation aus einer lokalen Relation lässt sich gut an folgenden Beispielen mit der Präposition *vor* illustrieren:

*Ich werde vor dem Kino stehen.
Er wird vor Einbruch der Dunkelheit eintreffen.*

Die Präposition *vor* im Satz *Ich werde vor dem Kino stehen* drückt eine lokale Relation aus. Im Satz *Er wird vor Einbruch der Dunkelheit eintreffen* zeigt dieselbe Präposition eine temporale Relation an, die sich aus der lokalen ableiten lässt: Die Präpositionen *vor* referiert in lokaler Verwendung auf die Stellung relativ zu einem Objekt bzw. zu einer Person, in temporaler Verwendung verweist sie auf die zeitliche Verortung eines Sachverhalts relativ zu einem anderen Sachverhalt: In beiden Fällen ist eine Vor-Orientierung gemeint, einmal eine lokale und einmal eine temporale Vor-Orientierung (vgl. Eisenberg 2013: 186).

Die Klassifizierung spezifischer präpositionaler Relationen als lokal, temporal, kausal oder modal reflektiert den vorhandenen semantischen Gehalt dieser Präposition: Eben weil die Präposition einen semantischen Gehalt aufweist, kann diese näher klassifiziert und die Relation semantisch beschrieben werden. Im Gegensatz dazu lässt sich in gewissen syntaktischen Kontexten von Präpositionen bzw. Präpositionalphrasen eine semantische Relation jedoch nicht (eindeutig) bestimmen. Präpositionen drücken dann ein Verhältnis aus, das Duden (2016: 618) als „neutral“ bezeichnet. Die neutrale Relation steht dabei für solche

Verwendungsweisen von Präpositionen, in denen Präpositionen ihre „eigentliche Bedeutung“, die sie in anderen syntaktischen Kontexten ausdrücken, nicht „entfalten“ (Duden 2016: 618). Solche Präpositionen werden z. B. von einem Verb oder einem Adjektiv regiert und sind i.d.R. nicht austauschbar. Ich spreche an dieser Stelle von *Präpositionsreaktion* (in Abgrenzung zur *Kasusreaktion*) und meine damit die Präpositionen, die von Verben oder Adjektiven regiert werden. Hierzu zählen ausschließlich primäre Präpositionen (Duden 2016: 618, Breindl 1989: 10). Die Präpositionen gelten dann als „bedeutungsleer oder bedeutungsneutral“ (Duden 2016: 618).¹² Im Gegensatz dazu spreche ich von einer *autonomen Präposition*, wenn keine Präpositionsreaktion vorliegt. Beispiele für Präpositionsreaktion sind:

Sie warteten auf den Schulbus.

Er verliebte sich in seine Nachbarin.

Die Relation, die neutrale Präpositionen ausdrücken, ist semantisch nicht (näher) beschreibbar. Die Präposition ist vielmehr ein „Verbindungsglied ohne eigene Bedeutung“ (Duden 2016: 618), sie ist i. d. R. nicht austauschbar und verfügt lediglich über eine grammatische Funktion.

Hinsichtlich der Unterscheidung von Präpositionen mit semantischer Bedeutung und neutraler Bedeutung spricht Duden (2016: 614) anderenorts ebenfalls vom *freiem Gebrauch* im Unterschied zum *gebundenem Gebrauch* von Präpositionen und stellt damit die Austauschbarkeit von Präpositionen in den Vordergrund. Beim freien Gebrauch trägt die Präposition eine eigenständige Bedeutung und kann durch andere Präpositionen ausgetauscht werden. Beim gebundenen Gebrauch dahingegen ist die Präposition bedeutungsneutral und kann nicht durch eine andere Präposition ersetzt werden¹³. Beispiele für freien und gebundenen Gebrauch sind im Folgenden aufgeführt. Bei diesen Beispielen handelt es sich ebenfalls um Beispiele für autonome Präpositionen (freier Gebrauch) bzw. für Präpositionsreaktion (gebundener Gebrauch):

Freier Gebrauch:

Das Buch liegt auf/unter/neben/hinter dem Tisch.

Gebundener Gebrauch

*Der Gast wartet auf/*um das Essen.*

¹² Nach Rostila (2018) kann die Mehrheit regierter Präpositionen bzw. die entsprechenden PP sehr wohl semantisch beschrieben werden.

¹³ Gebundener Gebrauch tritt nach Duden (2016: 614f.) ebenfalls bei Funktionsverbgefügen sowie in komplexen Adverbialien auf. Duden (2016: 614f.) nennt hier z. B. *am Tag, in der Nacht* oder *um Mitternacht*. Leider gehen die AutorInnen nicht weiter darauf ein, aus welchen Gründen der Gebrauch von Präpositionen in komplexen Adverbialien als *gebundener Gebrauch* gilt.

Nahezu alle primären Präpositionen können in Kontexten auftreten, in denen sie eine neutrale Relation ausdrücken (vgl. Duden 2016: 618). Folglich handelt es sich bei der neutralen Relation um ein Phänomen, das vor allem die hochfrequenten Vertreter der Wortart betrifft. Man hat es somit mit Präpositionen wie *an* oder *auf* zu tun, die in entsprechenden syntaktischen Kontexten entweder eine Bedeutung aufweisen oder bedeutungsneutral sind (s. obige Beispiele).

In der Darstellung bedeutungsneutraler Präpositionen im Duden (2016: 618) wird die Bedeutungsneutralität implizit mit der Präpositionsreaktion des Verbs (Adjektivs o. Substantivs) in Verbindung gesetzt. Eine bedeutungsneutrale Präposition ist demnach eine Präposition, die (von einem Verb oder einem Adjektiv) regiert wird (vgl. Duden 2016: 614). An dieser Stelle sei jedoch bereits einschränkend erwähnt, dass die Präpositionsreaktion nicht ausschließlich bedeutungsneutrale Präpositionen umfasst (vgl. Breindl 1989: 42). Ich gehe weiter unten im Kontext der syntaktischen Funktionen von Präpositionalphrasen näher auf diesen Punkt ein.

Die obigen Ausführungen machen deutlich, dass sich Präpositionen hinsichtlich des Verhältnisses, das sie ausdrücken können, in zwei Gruppen einteilen lassen:

- 1) Präpositionen, die eine semantisch beschreibbare Relation ausdrücken und damit in ihrer Bedeutung gut isolierbar sind. Diese Präpositionen nenne ich *semantisierte Präpositionen*. Es kann angenommen werden, dass semantisierte Präpositionen die prototypische Verwendung von Präpositionen in der deutschen Sprache widerspiegeln, d. h. eine prototypische Präpositionsverwendung drückt eine semantisch beschreibbare Relation aus.
- 2) Präpositionen, die eine semantisch nicht (bzw. nicht gut) beschreibbare Relation ausdrücken und damit nicht gut in ihrer Bedeutung isoliert werden können. Diese Präpositionen nenne ich *desemantisierte Präpositionen*. Präpositionen werden hier folglich hinsichtlich ihres semantischen Gehalts (semantisiert/desemantisiert) unterschieden. Tabelle 2 gibt hierzu einen zusammenfassenden Überblick.

Semantischer Gehalt der Präposition	
semantisiert	desemantisiert
Relation semantisch beschreibbar	Relation semantisch nicht bzw. nicht gut beschreibbar
Lokale Relation	Neutrale Relation
Temporale Relation	
Kausale Relation	
Modale Relation	

Tabelle 2: Semantischer Gehalt von Präpositionen: Überblick.

Der semantische Gehalt einer Präposition kann ausschließlich im konkreten syntaktischen Kontext der entsprechenden PP ermittelt werden. Aus diesem Grund gehe ich unter Kapitel 1.2.1 näher auf den semantischen Gehalt von Präpositionen ein.

1.1.3 Verschmelzung der Präposition mit definitivem Artikel

Einige Präpositionen können mit einem definiten Artikel¹⁴ im Dativ oder Akkusativ zu einer einzigen Wortform verschmelzen. So wird z. B. aus den Vollformen *in + dem*, *an + dem* und *bei + dem* die Verschmelzungen *im*, *am* und *beim*. Bei den Präpositionen, die mit dem definiten Artikel verschmelzen, handelt es sich ausschließlich um primäre Präpositionen. Viele Verschmelzungsformen (z. B. *im*, *am*, *zum*, *beim*) sind fest in der Schriftlichkeit etabliert, zum Teil kommen die verschmolzenen Formen sogar deutlich häufiger vor als die entsprechenden Vollformen (vgl. Duden 2016: 628). Eher seltener verschmelzen die Präpositionen *hinter*, *für*, *unter*, *durch*, *auf* und *um*. Aktuelle Tendenzen im Sprachwandel, so Duden (2016: 630), zeigen jedoch, dass die eher umgangssprachlichen Verschmelzungsformen mit dem Artikel *dem* und *das*, nämlich *vorn*, *hintern*, *übern*, *fürs*, *durchs* u. a. ebenfalls in der Schriftsprache vertreten sind. Dahingegen sind laut Duden (2016: 630) Verschmelzungen mit dem akkusativischen definiten Artikel *den* (*vorn*, *hintern*, *untern*, *fürn* u. a.) sehr selten.

In manchen Kontexten sind Verschmelzungsformen mit definiten Artikeln obligatorisch, die nicht verschmolzenen Vollformen gelten als ungrammatisch (z. B. **in das Stocken geraten*). In anderen Kontexten gibt es zwischen Verschmelzungs- und Vollformen funktionale Unterschiede (*Sie geht zum Arzt* vs. *Sie geht zu dem Arzt*) (Näheres s. Duden 2016: 628ff.).¹⁵

Der Gebrauch von Verschmelzungsformen aus Präposition und Artikel ist an verschiedene Bedingungen geknüpft und betrifft nicht alle Präpositionen des Deutschen. Beispielsweise sind die primären Präpositionen nicht alle gleichermaßen fähig, mit einem Artikel zu verschmelzen, obwohl durch aktuelle Sprachwandelprozesse sich weitere Verschmelzungsformen herausbilden (Duden 2016: 630).

Wichtig ist an dieser Stelle zu betonen, dass die Verschmelzungen aus Präposition und Artikel keinen Einfluss auf die die Bedeutung bzw. den semantischen Gehalt der beteiligten Präposition haben. Die Relation, die durch die Präposition ausgedrückt wird, wird durch das Vorhandensein einer Verschmelzungsform nicht verändert, vgl. *im Supermarkt* und *in dem Supermarkt*. In den empirischen Untersuchungen der vorliegenden Arbeit werden Verschmelzungsformen aus Präposition und Artikel bei der Auswertung berücksichtigt, es erfolgt jedoch keine weitere Analyse des Gebrauchs von Verschmelzungs- gegenüber Vollformen. Untersuchungen in Lernerkorpora zum Gebrauch von Verschmelzungs- und

¹⁴ In der gesprochenen Sprache sind zusätzlich Verschmelzungsformen mit indefinitem Artikel zu beobachten (*auf'nem*, *auf'ner*, *mit'ner* u. a.), wobei es bei den Verschmelzungen keine Restriktionen in Bezug auf die indefiniten Artikelformen gibt (Duden 2016: 631). Eisenberg (2013: 193) macht darauf aufmerksam, dass rein formal betrachtet nicht eindeutig ist, ob ein definiten oder ein indefiniten Artikel in einer Verschmelzungsform vorliegt: In der Verschmelzungsform *dem* könnte das *-m* sowohl für *einem* als auch für *dem* stehen.

¹⁵ Die Obligatorik von Verschmelzungsformen (auch in der Schriftlichkeit) wird allerdings nicht immer als solche wahrgenommen. So verweist Breindl (2013: 6) im Rahmen ihrer Untersuchung zum Gebrauch von Verschmelzungsformen auf die vertretene Meinung (von Muttersprachlern und Nicht-Muttersprachlern), Verschmelzungsformen aus Präposition und Artikel seien generell als umgangssprachlich einzustufen. Aus diesem Grund können Hyperkorrekturen vorliegen, wenn Artikel und Präposition in schriftlichen Texten in ihrer nicht verschmolzenen Form auftreten, obwohl der Kontext eine obligatorische Verschmelzungsform fordert.

Vollformen durch Deutschlernende liegen mit Breindl (2013) und Sieradz/Bordag (2014) vor (s. 4.2.2).

Die Ausführungen dieses Kapitels machen deutlich, dass Präpositionen in der deutschen Sprache ohne Berücksichtigung ihres Gebrauchskontexts nicht adäquat beschrieben werden können. Vor allem im Hinblick auf den semantischen Gehalt der Präposition spielt die größere Einheit der Präpositionalphrase eine entscheidende Rolle.

1.1.4 Zusammenfassung der zentralen Terminologie I

Tabelle 3 fasst die zentralen Termini des vorangegangenen Kapitels in einem Überblick zusammen.

Term	Erläuterung
Primäre Präposition	<p>Diese Präpositionen gehören zur ältesten Schicht und damit zum Kernbestand der deutschen Präpositionen. Primäre Präpositionen stellen die prototypischen Vertreter der Kategorie dar. Beispiele: <i>an, auf, mit, für</i>.</p> <p>Die primären Präpositionen, die hier behandelt werden, stehen vor ihrem Bezugselement.</p>
Bezugselement	<p>Die Einheit, die eine Präposition fordert. Beispiel: <i>über den Wolken</i>: Die Nominalphrase <i>den Wolken</i> ist das Bezugselement der Präposition <i>über</i>.</p>
Kasusreaktion	<p>Der/die Kasus, den/die eine Präposition fordert (regiert). Beispiel: <i>für</i> + Akkusativ: <i>Ich backe einen Kuchen für meinen Vater.</i> <i>aus</i> + Dativ: <i>Ich trage den Müll aus dem Haus.</i></p> <p>Wechselpräpositionen (z. B. <i>an, auf, in, unter</i>) können den Akkusativ und den Dativ regieren. Die Kasuswahl ist syntaktisch-semantic bestimmt.</p> <p>Die Kasusreaktion von Präpositionen wird in den empirischen Untersuchungen dieser Arbeit nicht untersucht.</p>
Regierte Präposition (Präpositionsreaktion) und autonome Präposition	<p>Die Präposition, die z. B. durch ein Verb oder ein Adjektiv regiert wird und daher i. d. R. nicht austauschbar ist. Beispiel: <i>Der Gast wartet auf/*um das Essen.</i></p> <p>Im Fokus dieser Dissertation steht die Präpositionsreaktion bei Verben und Adjektiven. Verben und/oder Adjektive regieren ausschließlich primäre Präpositionen (Duden 2016: 618, Breindl 1989: 10).</p>

	Bei autonomen Präpositionen liegt keine Präpositionsreaktion vor. Beispiel: <i>Das Buch liegt auf/unter/neben/hinter dem Tisch.</i>
Verschmelzung und Vollform	<p>Verschmelzung: Verschmelzung aus Präposition und definitivem Artikel (<i>am, im, beim</i>)</p> <p>Vollform: Die analytische Form (<i>an dem, in dem, bei dem</i>) der Verschmelzung.</p> <p>Verschmelzungen werden in der Untersuchung berücksichtigt, jedoch wird nicht zwischen dem Gebrauch von Verschmelzungen und Vollformen differenziert.</p>
Semantischer Gehalt von Präpositionen	<p>Bedeutungshaltigkeit der Präposition innerhalb einer Präpositionalphrase mit einer bestimmten syntaktischen Funktion.</p> <p>Semantisierte Präposition: Präposition, die eine semantisch beschreibbare Relation ausdrückt und in ihrer Bedeutung gut isolierbar ist.</p> <p>Desemantisierte Präposition: Präposition, die eine semantisch nicht (bzw. nicht gut) beschreibbare Relation ausdrückt und damit nicht gut in ihrer Bedeutung isoliert werden kann.</p>

Tabelle 3: Präpositionen: Zentrale Terminologie.

1.2 Präpositionalphrasen (PP)

Duden (2016: 782) definiert *Phrasen* als komplexe Einheiten, die aus syntaktischen Wörtern bestehen. Phrasen haben genau eine Wortform als Kern. Der Kern einer Phrase bzw. seine grammatischen Merkmale „bestimmen die syntaktischen Eigenschaften der daraus gebildeten Phrase in vielerlei Hinsicht“ (Duden 2016: 782). Mit grammatischen Merkmalen von Phrasenkernen sind hier die grammatischen Merkmale der entsprechenden Wortarten gemeint. Einzelne Phrasentypen im Deutschen werden auf Basis der Wortart ihrer Kerne klassifiziert.

Präpositionen fordern ein Bezugselement und regieren den Kasus dieses Bezugselements. Damit legt die Präposition die syntaktischen Eigenschaften des Bezugselements fest. Mit diesem Bezugselement bildet die Präposition eine Phrase – die Präpositionalphrase (PP). Eine PP ist eine komplexe Einheit aus Präposition als Kern und einer eingebetteten Phrase, bei der es sich i.d.R. um eine Nominalphrase (NP) handelt (Duden 2016: 849). Im Satz *Das Buch lag auf dem Tisch* besteht die PP *auf dem Tisch* aus der Präposition *auf* – dem Kern der Phrase – und der eingebetteten NP *dem Tisch* (vgl. Duden 2016: 783). Weiterhin kann eine PP eine Pronominalphrase, eine Adverbphrase oder eine Adjektivphrase enthalten (vgl. Duden 2016: 783, 849f.). In den folgenden Beispielen sind die PP und die jeweils eingebetteten Phrasen in eckigen Klammern hervorgehoben.

Pronominalphrase: *Die Blumen sind [für [dich]].*

Adverbphrase: *[Seit [gestern]] gelten die neuen Tarife.*

Adjektivphrase: *Ich halte Ottos Vorschlag [für [sehr riskant]].*

Die eingebettete Phrase kann unterschiedlich komplex sein und ggf. selbst eine weitere eingebettete Phrase enthalten (vgl. die Adjektivphrase *sehr riskant*, die die eingebettete Adverbphrase *sehr* enthält). Einige PP erlauben auch modifizierende Ausdrücke zur Erweiterung, vgl. *mitten* im folgenden Beispiel:

[[Mitten] im [Zimmer]] steht ein Tisch.

Innerhalb von PP können ebenfalls Verschachtelungen von PP auftreten. Im folgenden Beispiel liegen z. B. zwei ineinander verschachtelte PP vor (vgl. Duden 2016: 850):

Dieses Gerät stammt [von [vor [dem Krieg]]].

Im obigen Beispiel bildet die NP *dem Krieg* mit der Präposition *vor* die PP *vor dem Krieg*, die wiederum gemeinsam mit der Präposition *von* die verschachtelte PP *von vor dem Krieg* bildet. Zu Verschachtelungen von PP siehe z. B. Duden (2016: 850).

In diesem Kapitel werden PP aus unterschiedlichen Perspektiven beleuchtet. Zunächst werden die syntaktischen Funktionen von PP behandelt (Kap. 1.2.1). Hierunter wird die „inhaltliche Leistung, [der] semantische Beitrag zum Satz“ durch die PP verstanden (Duden 2016: 792). Hier stehen vor allem die Funktionen Objekt und Adverbiale im Vordergrund. Daraufhin folgt unter valenzgrammatischer Perspektive die Unterscheidung von PP als Ergänzungen bzw. als Angaben sowie eine Erläuterung der semantischen Rollen, die PP kodieren können (Kap. 1.2.3). Abschließend werden weitere Verwendungsweisen von PP dargelegt (u. a. Funktionsverbgefüge und *am*-Progressiv) (Kap. 1.2.5)¹⁶. Das Kapitel zu PP schließt mit einer sprachkontrastiven Perspektive auf Präpositionen und PP (Kap. 1.3). Hierbei steht der Vergleich von ausgewählten, dem Deutschen typologisch ähnlichen Sprachen mit ausgewählten, vom Deutschen typologisch verschiedenen Sprachen im Vordergrund. Im Anschluss an die einzelnen Kapitel wird jeweils die zentrale Terminologie zusammengefasst (Kap. 1.2.2, 1.2.4). Kapitel 1.4 bietet eine Zusammenfassung.

1.2.1 PP unter syntaktischer Perspektive: Syntaktische Funktionen

PP können unterschiedliche syntaktische Funktionen einnehmen. Duden (2016: 851) unterscheidet zwischen Präpositionalobjekt, Prädikativ, Adverbiale und Attribut. Präpositionalobjekt, Prädikativ und Adverbiale stellen Satzglieder dar, das Attribut repräsentiert eine Phrase im Inneren eines Satzglieds und stellt damit ein Gliedteil dar (Duden 2016: 806). An dieser Stelle sei bereits angemerkt, dass in der vorliegenden Arbeit die

¹⁶ In diesem Kapitel werden ausschließlich PP mit Präpositionen behandelt. Zur Struktur von PP mit Post- und Zirkumpositionen siehe z. B. Duden (2016: 849) oder GDS (1997: 2086f.).

syntaktischen Funktionen Präpositionalobjekt und Prädikativ aus forschungspragmatischen Gründen unter dem Begriff Objekt-PP zusammengefasst werden (Näheres s. unten). Die folgende Tabelle (Tab. 4) fasst die einzelnen Funktionen von PP, so wie sie für die vorliegende Arbeit angesetzt werden, in einem Überblick zusammen. Die PP sind jeweils unterstrichen.

Syntaktische Funktion		Beispiel
Objekt-PP	Präpositionalobjekt	<i>Julia interessiert sich <u>für Vulkanologie</u>.</i> <i>Anna ist stolz <u>auf ihre Tochter</u></i>
	Prädikativ	<i>Der Psychiater erklärte den Patienten <u>für völlig normal</u>.</i>
Adverbiale PP		<i>Anna reist <u>nach Portugal</u>.</i> <i><u>Seit dem Essen</u> sind vier Stunden vergangen.</i>
Attributive PP (Gliedteil)		<i>Der Weg <u>auf die Burg</u> war sehr steil.</i>

Tabelle 4: Syntaktische Funktionen von PP: Überblick

Die syntaktischen Funktionen können neben PP ebenfalls mit anderen strukturellen Einheiten realisiert werden. Beispiel: *Ich warte [darauf]*. Hier wird die syntaktische Funktion des Präpositionalobjekts als Adverbphrase (Präpositionaladverb) realisiert (vgl. Duden 2016: 794). Im Folgenden werden ausschließlich die syntaktischen Funktionen von PP behandelt, da diese für die vorliegende Arbeit zentral sind.

1.2.1.1 Adverbiale PP

Adverbiale PP dienen, grob gesagt, der Situierung eines Sachverhalts (Eisenberg 2013: 183). Duden (2016: 792ff.) gibt keine Definition für Adverbialien, jedoch wird in den entsprechenden Ausführungen (Duden 2016: 794ff.) deutlich, dass Adverbialien modifizierend bzw. situierend wirken. Bei den Adverbialien unterscheidet Duden (2016: 794) zwischen zwei Hauptverwendungsweisen: *Kommentaradverbialien* und *Situativadverbialien*, vgl. die Beispiele unten. Im Folgenden gehe ich ausschließlich auf Adverbialien ein, die durch PP realisiert werden (adverbiale PP).

Kommentaradverbiale: *Für mein Gefühl ist der Tee zu bitter.*

Situativadverbiale: *Mona arbeitet in Frankfurt.*

Kommentaradverbialien beziehen sich auf die ganze Äußerung und dienen der/m Sprecher/-in zur Beurteilung dieser bzw. ermöglichen die Stellungnahme zu einer Äußerung. Im oberen Beispiel wird durch die PP *für mein Gefühl* ausgedrückt, dass es sich um eine subjektive Meinungsäußerung handelt. Situativadverbialien dahingegen können sich auf das Prädikat oder auf den ganzen dargestellten Sachverhalt beziehen, vgl. das entsprechende Beispiel oben. Sie können sich also ebenfalls wie Kommentaradverbialien auf die Äußerung als Ganzes beziehen, dienen jedoch nicht deren Kommentierung, sondern deren situativer Einbettung. Während Kommentaradverbialien sich sozusagen von außen auf die Äußerung beziehen, beziehen sich Situativadverbialien von innen auf eine Äußerung (vgl. Duden 2016: 795).

Bei den Situativadverbialien wird zumeist eine weitere (grobe) Klassifikation vorgenommen, und zwar in Lokaladverbialien, Temporaladverbialien, Adverbialien der Art und Weise, und Kausaladverbialien im weiteren Sinne (Beispiele s. unten)¹⁷ – eine Klassifikation, die sich ebenso bei der Klassifikation der semantischen Relationen von Präpositionen findet (s. Kap. 1.1.2). Hier zeigt sich erneut: Präpositionen können nur unter Berücksichtigung ihres Gebrauchskontexts ausreichend beschrieben werden. Die Zuordnung der Präpositionen zu ihren semantischen Relationen ergibt sich aus den prototypischen Verwendungskontexten entsprechender adverbialer PP.

Lokaladverbiale: *Karl arbeitet in München.*

Temporaladverbiale: *Sie blieb für kurze Zeit.*

Adverbiale der Art und Weise: *Er isst den Fisch mit dem Messer.*

Kausaladverbiale (i.w.S.): *Im Wiederholungsfall wird er belangt.*

In der vorliegenden Arbeit werden Adverbialien als eine geschlossene Klasse behandelt, eine weitere semantische Aufteilung erfolgt nicht. Für das Forschungsanliegen steht die semantische Klassifikation der Adverbialien nicht im Vordergrund.

Adverbialien können in der Bedeutung des Verbs „vorangelegt“ (Duden 2016: 798) sein und somit vom Verb regiert werden. Im Satz *Otto begibt sich auf den Balkon* kann das Lokaladverbiale, das durch die PP *auf den Balkon* realisiert ist, nicht weggelassen werden (**Otto begibt sich*). Adverbialien können aber auch frei zu einem Satz hinzufügbare sein, vgl. *Otto raucht auf dem Balkon eine Zigarette*: Hier kann das lokale Adverbiale *auf dem Balkon* weggelassen werden (*Otto raucht eine Zigarette*), es ist nicht in der Bedeutung des Verbs vorangelegt (Näheres dazu s. Kap. 1.2.3 unten). Prototypische adverbiale PP werden „[ü]berwiegend“ (Duden 2016: 799) nicht vom Verb regiert und können in einem Satz relativ frei hinzugefügt werden. Für die Präposition innerhalb einer adverbialen PP gilt daher, dass i.d.R. keine Präpositionsreaktion vorliegt, d. h. die Präposition ist autonom und semantisiert¹⁸ (vgl. Duden 2016: 852). Die Präposition kann ebenfalls ausgetauscht werden, auch wenn die Möglichkeiten hier z. T. begrenzt sein können (Duden 2016: 852f., vgl. auch Breindl 1989: 29f.):

Der Wolf versteckt sich in/vor/hinter/an/auf dem Schrank.

Anna reiste mit/ohne Gepäck.

Bei einigen adverbialen PP kann man jedoch beobachten, dass die in der PP enthaltene Präposition nicht austauschbar ist (Breindl 1989: 32). Dies ist z. B. beim Verb *aussteigen* der Fall. Zur Referenz auf das Transportmittel als Ausgangsort kann lediglich eine PP mit *aus*

¹⁷ Einige Subklassen sind durch bestimmte Präpositionen gekennzeichnet: Instrumentale adverbiale PP enthalten beispielsweise nur die Präpositionen *mit* und *durch* (Breindl 1989: 32).

¹⁸ Der semantische Gehalt einer Präposition in adverbialen PP hat für Helbig/Buscha (2001a: 184) einen zentralen Stellenwert. Sie gehen davon aus, dass „[e]ine eigene Semantik [...] die Präpositionen im Wesentlichen nur [haben], wenn sie zum Ausdruck adverbialer Verhältnisse dienen“.

verwendet werden: *Ich steige aus dem Zug aus* (vgl. E-Valbu¹⁹), in solchen Fällen liegt dann bei adverbialen PP Präpositionsrektion vor; die Präposition ist jedoch semantisiert.

Adverbiale PP enthalten jedoch nicht immer semantisierte Präpositionen. Wie weiter oben bereits erwähnt, merkt Duden (2016: 615) an, dass desemantisierte Präpositionen ebenfalls in komplexen Adverbialien vorkommen können. Duden (2016: 615) nennt hier die PP *am Tag, in der Nacht, um Mitternacht*, aus denen hervorgeht, dass Tageszeitangaben in Form von PP unterschiedliche Präpositionen enthalten können, auch wenn sie allesamt ähnliche semantische Relationen ausdrücken (vgl. dazu auch Helbig/Buscha 2001b: 362ff.). Hier wird die Präpositionswahl durch die Nominalphrase (NP) in der PP bestimmt, sodass ein Austausch der NP mit einem Austausch der Präposition einhergeht (vgl. Breindl 1989: 30):

Martin kommt an diesem Tag/in diesem Monat/zu dieser Zeit.

Die Verwendung der Präposition ist an die Länge der zeitlichen Intervalle und damit an die mentale Konzeptualisierung dieser Intervalle geknüpft, sie kann in solchen Fällen daher als konventionalisiert gelten. Die Präposition ist zwar autonom, da keine Präpositionsrektion vorliegt, sie ist jedoch nicht im gleichen Maße semantisiert wie in prototypischen adverbialen PP.

Die syntaktische Funktion Adverbiale repräsentiert die zentrale Funktion einer PP, gleichzeitig wird ein Adverbiale i.d.R. in Form einer PP realisiert (Breindl 1989: 10). D. h. die Form PP ist prototypisch mit der Funktion Adverbiale verbunden und umgekehrt. Damit liegt hier eine recht zuverlässige Form-Funktionsverbindung vor.

1.2.1.2 Objekt-PP

Unter Objekt-PP werden in der vorliegenden Arbeit die syntaktischen Funktionen Präpositionalobjekt und Prädikativ zusammengefasst (Objekt-PP). Im Folgenden soll kurz auf die einzelnen Funktionen eingegangen und die Zusammenfassung beider Funktionen motiviert werden.

Unter einem Präpositionalobjekt versteht Duden (2016) eine PP in der Funktion eines Objekts. Dabei repräsentiert die PP in dieser Funktion die semantische Funktion eines „Mitspieler[s]“ oder „Rollenträger[s]“ bei einem Vorgang oder einer Handlung“ (Duden 2016: 792). Die PP drückt somit „die von einer Handlung betroffene Person oder Sache aus“ (Duden 106: 851). Die entsprechende PP ist abhängig von einem Verb, es liegt Präpositionsrektion vor. Im folgenden Satz regiert das Verb (*sich*) *kümmern* eine PP mit der Präposition *um*:

Petra kümmert sich um das Auto.

Die Präposition ist obligatorisch und fest, d. h. sie kann in der Regel nicht durch eine andere Präposition ausgetauscht werden (**Petra kümmert sich auf das Auto*) (vgl. auch Breindl 2008: 946, Helbig/Buscha 2001a: 184). In manchen Fällen scheint die Präposition in Präpositionalobjekten austauschbar zu sein (vgl. Beispiel unten). Jedoch liegen hier zwei

¹⁹ <https://grammis.ids-mannheim.de/verbs/view/400342/1> (März 2019).

verschiedene Verblesarten vor, die dann je eine andere Präpositionsreaktion aufweisen. Der Austausch der Präposition bedingt eine andere Lesart des Verbs (Breindl 1989: 31). Die Konstruktion *klagen um* meint *seine Trauer über den Verlust von etwas äußern*, während *klagen über* die Lesart *sich unzufrieden über etwas äußern* impliziert.

Max klagt um seinen Vater

Max klagt über seinen Vater

In manchen Fällen bewirkt ein Austausch der Präposition in Präpositionalobjekten dahingegen keine fassbare Veränderung in der Verblesart (*Hubert spricht dauernd über/von Herbert*, Duden 2016: 853). Nach Breindl (1989: 31) handelt es sich hier um Varianten von PP.

Bei anderen Fällen kommutieren die Präpositionen *für* und *gegen* in Präpositionalobjekten, wobei keine Veränderung der Verblesart beobachtet werden kann: *Das spricht für/gegen ihn*. Nach Breindl (1989: 32) zeigen solche Präpositionalobjekte adverbialtypisches Verhalten und stellen daher einen Sonderfall innerhalb der Präpositionalobjekte dar.

Zu den Präpositionalobjekten zählen nach Duden (2016: 851) ebenfalls PP, die von Adjektiven (auch: komplexe Prädikate²⁰) regiert werden, z. B. *zufrieden (sein)* oder *stolz (sein)*, siehe dazu die folgenden Beispiele. Ich folge dieser Klassifikation und zähle PP wie die unteren ebenfalls zu den Objekt-PP:

Anna ist stolz auf ihre Tochter.

Die Experten waren mit den Ergebnissen zufrieden.

Bei PP, die von Adjektiven regiert werden, liegt ebenfalls Präpositionsreaktion vor, und die Präposition ist nicht austauschbar. Innerhalb einer prototypischen Objekt-PP ist die Präposition desemantisiert und trägt folglich nicht zur Gesamtbedeutung der PP bei. In unterschiedlichen Grammatiken wird eine desemantisierte Präposition als konstitutiv für Präpositionalobjekte angesehen (vgl. Eisenberg 2013: 3014, Helbig/Buscha 2001a: 184, Grieshaber 2007a: 630). Allerdings gibt es hier auch Ausnahmen. Für die Präpositionen *mit* und *für* (vgl. die Beispiele unten) wird beispielsweise angenommen, dass diese in Objekt-PP semantisiert sind (vgl. Breindl 2006: 946, Breindl 1989: 39). Auf diesen Punkt komme ich weiter unten nochmal zurück.

Hans stimmt für Partei X.

Peter trifft sich mit Claudia.

²⁰ Nach Duden (2016: 866f.) werden u. a. Verbindungen aus Kopulaverb und einer adjektivisch prädikativen Ergänzung als komplexe Prädikate angesehen, bei denen das Adjektiv eine starke Integration ins Prädikat aufweist (z. B. *zufrieden sein*). Eine ähnliche Auffassung findet sich bei Hoffmann (2016: 248).

Prädikative PP machen „eine Aussage über eine oder mehrere andere Phrasen“ (Duden 2016: 799). Diese andere Phrase, über die die PP eine Aussage macht, wird Bezugsphrase genannt. Die Bezugsphrase kann entweder Subjekt oder Objekt des Satzes sein, es kann somit Subjekt- oder Objektbezug einer prädikativ verwendeten PP vorliegen, vgl. 11. und 12. unten. Man spricht hier auch von einem Subjektsprädikativ (11.) bzw. einem Objektsprädikativ (12.). Zur semantischen Klassifizierung von Prädikativen s. Duden (2016: 801ff.).

11. *Sie wurde zur Präsidentin gewählt.*

12. *Der Psychiater erklärte den Patienten für völlig normal.*

Prädikative PP werden, wie Präpositionalobjekte, vom Verb regiert, es liegt Präpositionsrektion vor (vgl. Duden 2016: 852).

In dieser Arbeit werden Präpositionalobjekte und prädikative PP zusammengefasst und als *Objekt-PP* klassifiziert. Die Zusammenlegung ist u. a. dadurch motiviert, dass in beiden Funktionen Präpositionsrektion vorliegt. Gemein ist prädikativen PP und Präpositionsobjekten weiterhin, dass der semantische Gehalt der Präposition in beiden Funktionen als „meist gering, teilweise sogar ganz vernachlässigbar“ eingestuft wird (Duden 2016: 852), beide Funktionen enthalten somit i.d.R. desemantisierte Präpositionen²¹. Im Übrigen verweist Duden (2016: 852) ebenfalls auf die geläufige Zusammenlegung von Präpositionalobjekten und prädikativen PP in anderen Grammatiken²².

Innerhalb der Kasushierarchie wird der Objekt-PP ein besonderer Status zugeschrieben. Nach Primus (1999: 166) repräsentiert die Kasushierarchie einen Ansatz zur Erklärung einzelner Rektionsmuster im Gegenwartsdeutschen. Die Kasushierarchie, wie sie hier aufgefasst wird, ist nach Primus (1999: 135) im theoretischen Rahmen der Optimalitätstheorie zu verorten, die sich wiederum in das Paradigma der generativen Grammatik einbettet (Elsen 2014: 149). In der generativen Grammatik geht man davon aus, dass die Kasus Nominativ und Akkusativ strukturelle Kasus sind, Dativ und Genitiv gelten dahingegen als inhärente bzw. lexikalische Kasus (Philippi 2008: 125): Es wird angenommen, dass der Kasus Nominativ (im Regelfall) immer dem Subjekt zugewiesen wird und der Kasus Akkusativ immer dem (ersten) Objekt. Das Auftreten der lexikalischen Kasus Dativ und Genitiv muss dahingegen explizit „gelernt werden“, da ihr Auftreten von Einzelverben abhängt (Philippi 2008: 126). Implizit wird dadurch eine Hierarchie der Kasusformen²³ auf Basis der Positionen von Einheiten im Satz etabliert: Nominativ – Akkusativ – Dativ/Genitiv. Objekt-PP werden in diesem Zusammenhang bei Philippi (2008: 126) nicht thematisiert.

²¹ Breindl (2006) merkt jedoch an, dass u.a. die Präposition *für* (wie sie ja auch in prädikativen PP enthalten sein kann) „in keiner Verwendung desemantisiert“ ist (946).

²² Breindl (1989: 12f.) differenziert zwischen Präpositionalobjekten und prädikativen PP und schließt letztere aus dem Gegenstandsbereich ihrer Arbeit zu Präpositionalobjekten aus. In ihren kurzen Ausführungen zu prädikativen PP merkt Breindl (1989: 13) jedoch an, dass einige prädikative PP Eigenschaften mit Präpositionalobjekten teilen.

²³ Die Ergebnisse einer Untersuchung zum Kasuserwerb durch Kinder mit Deutsch als Zweitsprache von Wegener (1995a: 352) stützen die Annahme dieser generativen Kasushierarchie und bestätigen eine entsprechende Rangfolge im Erwerb der Kasus.

Die Kasushierarchie (hier illustriert nach Eisenberg 2013: 63, vgl. Abb. 1) spiegelt die Markiertheitsverhältnisse von Kasusformen und Kasusfunktionen wider, die linke Seite repräsentiert dabei die hierarchisch höchste Stelle, die rechte Seite die hierarchisch niedrigste. Markiertheit meint in diesem Zusammenhang die formale Kodierung der einzelnen Formen sowie die Distribution der durch die einzelnen Formen realisierten syntaktischen Funktionen in der Sprache. Damit zeigt die Kasushierarchie an, dass die einzelnen Kasusformen „nicht gleichwertig“ sind (Duden 2016: 817)²⁴.



Abbildung 1: Kasushierarchie im Deutschen (nach Eisenberg 2013: 63).

Der Nominativ (bzw. das Subjekt) repräsentiert die ranghöchste Stellung (Abb. 1). Eisenberg (2013: 63) siedelt den Genitiv (bzw. das Genitivobjekt) und die PP (bzw. Objekt-PP) auf der gleichen Stufe an. Er merkt an, dass diese „Gleichstellung“ dadurch motiviert ist, dass beide Formen „gegenüber den anderen Kasus als hierarchisch niedriger einzustufen [sind]“ (Eisenberg 2013: 64). Primus (1999: 139) weist darauf hin, dass es schwierig ist, den Status einer PP bzw. Objekt-PP innerhalb der Kasushierarchie genau zu bestimmen. Sie ordnet PP bzw. Objekt-PP jedoch als hierarchisch niedriger ein als den Dativ bzw. das Dativobjekt (Primus 1999: 164).

In der Kodierung einer PP muss sowohl die Präpositionsreaktion als auch die Kasusreaktion beachtet werden. Die formale Realisierung einer PP erfordert demnach Kenntnisse zur Präpositions- und zur Kasusreaktion. Verglichen mit z. B. dem Nominativ oder dem Akkusativ zeigt die PP somit einen höheren (formalen) Kodierungsaufwand. Gleichzeitig spiegelt die Kasushierarchie die Präferenz von Kasusformen wider und verweist auf eine Hierarchie der einzelnen syntaktischen Funktionen (Eisenberg 2013: 64): Der Nominativ gilt als „Normalkasus“ (Duden 2016: 817), die syntaktische Funktion *Subjekt* (die durch den Nominativ realisiert wird) ist „wie kein anderes Satzglied in syntaktische Regularitäten involviert“ (Eisenberg 2013: 64); es gibt mehr Verben, die ein Akkusativobjekt regieren als Verben, die ein Dativobjekt regieren und noch weniger Verben regieren Objekt-PP (vgl. Eisenberg 2013: 64). Verglichen mit Akkusativ- oder Dativobjekten sind Objekt-PP folglich seltener im Sprachsystem vertreten. Gemäß der Kasushierarchie (Abb. 1) nimmt die Objekt-PP demnach insgesamt eine besondere (Rand-)Stellung ein. Die Objekt-PP stellt hierarchisch betrachtet an letzter Stelle und ist damit hierarchisch niedriger einzustufen als z. B. ein Akkusativ- oder Dativobjekt. Die Kasushierarchie macht deutlich, dass Objekt-PP a) einen vergleichsweise hohen Kodierungsaufwand haben und b) weniger frequent sind als reine

²⁴ Die empirische Grundlage der Kasushierarchie wird weder in Eisenberg (2013: 63 ff.) noch in Duden (2016: 817) explizit gemacht. In Bezug auf die Häufigkeit gewisser Kasusformen verweist Eisenberg (2013: 64) an einer Stelle auf Primus (1999: 139), die auf Auszählungen auf Grundlage des Verbwortschatzes in Mater (1971) verweist. Duden (2016: 817) nennt in Zusammenhang mit der Kasushierarchie bzw. Rangfolge lediglich formale Aspekte (z. B. Endungen).

Kasusobjekte. Daraus kann geschlussfolgert werden, dass eine Objekt-PP keine prototypische Form von Verbobjekten repräsentiert.

Zusammenfassend möchte ich an dieser Stelle festhalten, dass prototypische Objekt-PP und adverbiale PP Unterschiede hinsichtlich ihrer Form-Funktionsverbindungen aufweisen. Unter Form verstehe ich hier die strukturelle Einheit der PP, als Funktion zählt hier die syntaktische Funktion. Adverbiale PP repräsentieren prototypische Form-Funktionsverbindungen, da Adverbialien i.d.R. in der Form von PP realisiert werden (Heidolph/Flämig/Mosch 1981: 370) und PP in der Regel in der Funktion von Adverbialien auftreten. Die entsprechende Form-Funktionsverbindung ist somit zuverlässig. Dahingegen wird anhand der Kasushierarchie deutlich, dass die syntaktische Funktion Objekt prototypisch in Form eines Nominals realisiert wird, und vergleichsweise selten in Form einer PP. Diese Form-Funktionsverbindung ist somit weniger zuverlässig. Tabelle 5 fasst die unterschiedlichen Form-Funktionsverbindungen von Objekt-PP und adverbialer PP zusammen.

Form: Struktur	Funktion: Syntaktische Funktion	Form-Funktionsverbindung
PP	Adverbiale PP	prototypisch und zuverlässig
	Objekt-PP	nicht prototypisch und weniger zuverlässig

Tabelle 5: Objekt-PP und adverbiale PP: Form-Funktionsverbindungen.

Semantischer Gehalt der Präposition: Objekt-PP vs. adverbiale PP

Innerhalb einer prototypischen adverbialen PP ist die Präposition autonom und semantisiert und damit an der Konstruktion der Gesamtbedeutung der PP beteiligt (vgl. *Sie isst einen Apfel an der Straßenecke*). Damit steht die prototypische adverbiale PP im Kontrast zur prototypischen Objekt-PP. In einer prototypischen Objekt-PP hat die Präposition ihre ursprüngliche Bedeutung verloren, die Präposition ist desemantisiert. Somit trägt die Präposition nicht zur Bedeutung der Gesamt-PP bei (vgl. *Sie denkt an sein Versprechen*).

Betrachtet man die PP im Satz *Sie isst einen Apfel an der Straßenecke*, so kann man feststellen, dass die Präposition semantisiert ist und eine semantisch beschreibbare Relation ausdrückt. Die Bedeutung der Präposition *an* innerhalb der adverbialen PP *an der Straßenecke* ist gut isolierbar. Für die Bedeutung der PP *an der Straßenecke* gilt hier, dass diese verschieden ist von der Bedeutung der eingebetteten Nominalphrase (NP) (*der Straßenecke*) (vgl. GDS 1997: 2155). Dies bedeutet, dass die Bedeutung der Gesamt-PP sich zusammensetzt aus der Bedeutung der NP *die Straßenecke* und der Bedeutung der Präposition *an*. Aus der Kombination der Einheiten, bzw. der Bedeutung dieser Einheiten, kann die Bedeutung der PP *an der Straßenecke* abgeleitet werden.

Diese Art der kompositionalen Bedeutung trifft auf prototypische Objekt-PP nicht zu. Die Bedeutung der Objekt-PP *an sein Versprechen* im Satz *Sie denkt an sein Versprechen* ist hier gleichzusetzen mit der Bedeutung der eingebetteten NP (*sein Versprechen*) (vgl. GDS 1997: 2155). Es liegt Präpositionsrektion vor und die Bedeutung der Präposition *an* ist hier nicht gut isolierbar. Zur Bedeutung der PP *an sein Versprechen* trägt die Präposition demnach

nicht bei (vgl. GDS 1997: 2155). Abbildung 2 illustriert den Unterschied in der kompositionalen Bedeutung innerhalb prototypischer adverbialer PP und prototypischer Objekt-PP.

Adverbiale PP		
<i>Beispielsatz</i>	<i>Sie isst einen Apfel <u>an der Straßenecke</u></i>	
<i>Struktur</i>	Verb	PP [an [der Straßenecke]]
	essen	an der Straßenecke
<i>Bedeutung</i>	<i>etwas zu sich nehmen</i>	<i>P + NP = an + der Straßenecke</i>
Objekt-PP		
<i>Beispielsatz</i>	<i>Sie denkt <u>an sein Versprechen</u></i>	
<i>Struktur</i>	Verb	PP [an [sein Versprechen]]
	denken	an sein Versprechen
<i>Bedeutung</i>	<i>seine Überlegungen auf etwas richten</i>	<i>NP = sein Versprechen</i>

Abbildung 2: Kompositionale Bedeutung prototypischer adverbialer PP und Objekt-PP²⁵.

In Bezug auf Objekt-PP wie in Abbildung 2 kann man an dieser Stelle auch davon sprechen, dass der Inhalt, der über die Präposition transportiert wird, vom Verbinhalt präsupponiert wird (Lerot 1982: 287). Lerot (1982) spricht Präpositionen in Objekt-PP zwar keine inhaltliche Bedeutung ab, attestiert ihnen jedoch aus pragmatischer Perspektive „einen sehr geringen Informationsgehalt“ (Lerot 1982: 288). Nach Lerot (1982: 288) informieren regierte Präpositionen in Objekt-PP „den Hörer über Relationen, die ihm sehr gut bekannt sind, da sie zu seinem Vorwissen gehören“, aus pragmatischer Perspektive erscheinen solche Präpositionen damit „bedeutungslos“ (Lerot 1982: 288).

Die obigen Ausführungen bauen eine Dichotomie des semantischen Gehalts von Präpositionen innerhalb adverbialer PP und Objekt-PP auf: Adverbiale PP enthalten (immer) semantisierte Präpositionen, deren Bedeutung gut isolierbar ist, Objekt-PP dahingegen zeichnen sich (immer) durch desemantisierte Präpositionen aus, deren Bedeutung nicht gut isolierbar ist. Diese strikte Trennung in Bezug auf den semantischen Gehalt der Präposition in Objekt-PP und adverbialer PP gilt jedoch nicht in absoluter Weise. So weist z. B. Engelberg (2010: 115) darauf hin, dass der fehlende Bedeutungsbeitrag bei Präpositionen in Objekt-PP „nur in besonders typischen Fällen“ zu beobachten ist und sich oft „zumindest bestimmte Affinitäten zwischen der Form von Konstituenten und ihren relational-semantischen Bezug zur Verbbedeutung fest[...]stellen“ lassen (Engelberg 2010: 115). Im Folgenden gehe ich näher auf den semantischen Gehalt von Präpositionen in Objekt-PP ein.

Für Präpositionen in Objekt-PP lässt sich nach Breindl (1989) ein Bedeutungsspektrum aufstellen, das entsprechend Abbildung 3 veranschaulicht werden kann.

²⁵ Die jeweilige Verblesart ist E-Valbu entnommen (Stand: Mai 2019). Weitere Informationen zu E-Valbu s. weiter unten.

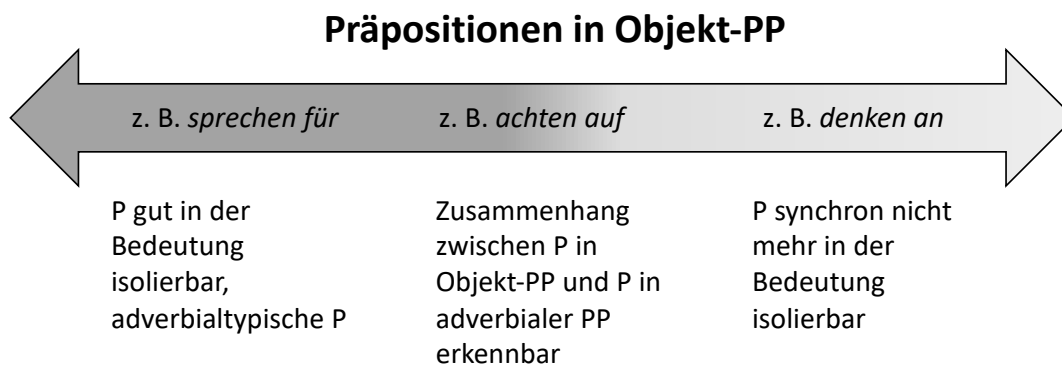


Abbildung 3: Bedeutungsspektrum von Präpositionen (P) in Objekt-PP (nach Breindl 1989: 42).

Am rechten Rand des Spektrums von Präpositionen in Objekt-PP finden sich Präpositionen, deren Bedeutung nicht mehr isoliert werden kann, während sich am linken Rand des Spektrums Präpositionen finden, die eine gut isolierbare Bedeutung aufweisen, da sie der präpositionalen Bedeutung in prototypischen adverbialen PP ähneln. Breindl (1989: 39) hält fest, dass verschiedene Präpositionen innerhalb von Objekt-PP – Breindl (1989) spricht hier von PO-Präpositionen – unterschiedliche Grade semantischer Transparenz aufweisen, wobei sie zwei zentrale Gruppen von Präpositionen nennt:

Von den PO-Präpositionen haben generell eine Tendenz zur ‚Durchsichtigkeit‘ und relativ konstante Bedeutungszüge die Präpositionen *mit*, *für*, *gegen*, *nach* und *zu*. Die Präpositionen *an*, *auf*, *aus*, *in*, *über*, *um*, *vor* und *von* kommen dagegen häufiger in PO²⁶ vor, in denen eine Rückführung auf eine Grundbedeutung der Präpositionen nicht oder nur mit einer etwas gekünstelten Bedeutungshypothese möglich ist.

Innerhalb von Objekt-PP gibt es bei Präpositionen folglich spezifische Tendenzen im Hinblick auf ihren semantischen Gehalt. Breindl (1989: 39) trennt hier zwischen Präpositionen mit gut isolierbarer Bedeutung und solchen mit weniger gut isolierbarer Bedeutung. Demnach unterscheidet sie zwischen semantisierten und desemantisierten Präpositionen und stellt hier beispielhaft die semantisierten Präpositionen *mit*, *für*, *gegen*, *nach* und *zu* den desemantisierten Präpositionen *an*, *auf*, *aus*, *in*, *über*, *um*, *vor* und *von* gegenüber. Die folgende Abbildung (Abb. 4) illustriert diese Gegenüberstellung.

²⁶ Die Abkürzung PO steht für *Präpositionalobjekt*.

Präpositionen in Objekt-PP

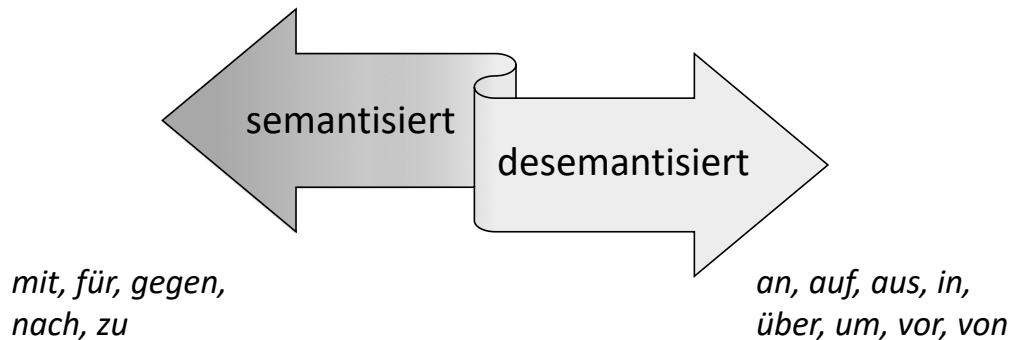


Abbildung 4: Semantisierte und desemantisierte Präpositionen in Objekt-PP (nach Breindl 1989: 39).

Im Folgenden werde ich anhand der desemantisierten Präpositionen *an* und *auf* sowie der semantisierten Präpositionen *mit* und *für* näher auf den unterschiedlichen semantischen Gehalt von Präpositionen in Objekt-PP eingehen.

Die Präpositionen *an/auf* gelten in Verwendung einer Objekt-PP als desemantisiert. Ihre prototypische Bedeutung, die die Präpositionen *an/auf* in lokalen adverbialen PP ausdrücken, verlieren sie in Objekt-PP²⁷. Die folgenden Beispiele mit *auf*-PP illustrieren diese Bedeutungsreduktion:

Adverbiale PP: *Helga sitzt auf der Bank*

Objekt-PP: *Helga wartet auf ihren Freund.*

Im obigen Beispiel illustriert die adverbiale PP die Grundbedeutung der Präposition *auf* bzw. einer PP mit der Präposition *auf*. Diese beschreibt eine Kontaktregion in der Vertikalen bzw. eine Stütze i.w.S. (GDS 1997: 2121). Die Objekt-PP kann dahingegen nicht mit der semantischen Relation einer ‚Stütze‘ in Beziehung gebracht werden. Die *auf*-PP weist hier „ein[en] Moment von Zielgerichtetheit/Prospektivität/Intentionalität“ (GDS 1997: 2121) auf. Ähnliche Verhältnisse in Bezug auf den semantischen Gehalt in Objekt-PP können auch für die Präposition *an* festgestellt werden. In den Beispielen unten kann die semantische Relation der adverbialen *an*-PP als Lokalisierung in einer Außenregion eines Objekts (GDS 1997: 2116f.), hier also der Wandfläche, umschrieben werden. Die semantische Relation der Objekt-PP mit *an* dahingegen kann als Referenz auf eine „kognitive Nähe“ umschrieben werden (GDS 1997: 2121).

Adverbiale PP: *Hans lehnt an der Wand*

Objekt-PP: *Hans denkt an seine Eltern.*

²⁷ Hier geht es mir nicht darum, ob die Präpositionen in adverbialen PP stets dieselbe semantische Relation ausdrücken bzw. wie viele verschiedene semantische Relationen sie realisieren können. Es geht hier primär darum, ob die Präpositionen in Objekt-PP i.d.R. einen semantischen Gehalt aufweisen oder nicht.

Wenn man nun davon ausgeht, dass die adverbiale PP i.d.R. die Grundbedeutung der Präposition ausdrückt, wird hier deutlich, dass im Fall der Objekt-PP mit *an/auf* eine Rückführung auf die Grundbedeutung nicht gut möglich ist (vgl. Zitat Breindl oben)²⁸. Diese Präpositionen können in Objekt-PP somit als desemantisiert angesehen werden.

Die Präpositionen *mit/für* können dahingegen in Objekt-PP als semantisiert klassifiziert werden. Die unteren Beispiele (entnommen aus Breindl 1989: 39) illustrieren anhand der Präposition *für*, dass semantisierte Präpositionen sowohl in adverbialen PP als auch in Objekt-PP einen ähnlichen, d. h. relativ konstanten semantischen Gehalt aufweisen. Der semantische Gehalt der Präposition ist hier noch gut zu erkennen bzw. sind die „gemeinsamen Bedeutungszüge von PO[Objekt-PP, TW] und adverbialer PP nicht zu übersehen“ (Breindl 1989: 39):

Adverbiale PP: *Hans arbeitet für Firma X.*

Objekt-PP: *Hans stimmt für Partei X.*

In den obigen zwei Beispielen referiert die *für*-PP unabhängig von der syntaktischen Funktion der PP jeweils auf den/die beabsichtigte Nutznießende/n der Handlung (vgl. GDS 1997: 2131). Hier repräsentieren die PP trotz unterschiedlicher syntaktischer Funktionen (Objekt-PP und adverbiale PP) identische semantische Relationen. Die Beispiele zeigen somit, dass die semantischen Relationen der *für*-PP in adverbialen PP sowie in Objekt-PP übereinstimmen können (vgl. GDS 1997: 2131). Ähnliches gilt ebenfalls für *mit*-PP in den folgenden Beispielen (nach Breindl 1989: 39):

Adverbiale PP: *Anna arbeitet mit Tom.*

Objekt-PP: *Anna unterhält sich mit Tom.*

Die Objekt-PP in reziprok-symmetrischer Verwendung steht der komitativen Verwendung der adverbialen *mit*-PP sehr nahe (vgl. GDS 1997: 2141ff). Somit zeigen die semantisierten Präpositionen *mit/für* im Hinblick auf den semantischen Gehalt der Präposition adverbialtypische Eigenschaften (1989: 269). Das bedeutet: Tendenziell liegt kein deutlicher Unterschied im semantischen Gehalt adverbialer PP und Objekt-PP mit *mit/für* vor.

Zusammenfassend kann festgehalten werden, dass bei *für*- und *mit*-PP ähnliche Bedeutungen der Präposition sowohl in Objekt-PP als auch in adverbialen PP vorliegen können, anders als bei *auf*- und *an*-PP. Die Präpositionen *an/auf* in adverbialen PP und Objekt-PP weisen i.d.R. keine konstanten Bedeutungszüge auf, ihre prototypische Bedeutung ist in Objekt-PP reduziert und die Präposition ist desemantisiert. Somit weisen diese Präpositionen eine weniger konstante bzw. weniger zuverlässige Verbindung zwischen ihrer Form und ihrer Bedeutung (bzw. Funktion) auf. Im Gegensatz dazu zeigen die Präpositionen *mit/für* in adverbialen PP und Objekt-PP i.d.R. ähnliche Bedeutungszüge, sodass ihre prototypische

²⁸ Rostila (2018) nimmt an, dass hierbei in vielen Fällen eine Bedeutungsveränderung der entsprechenden Präposition vorliegt.

Bedeutung in Objekt-PP nicht reduziert ist; diese Präpositionen sind semantisiert. Damit weisen die Präpositionen *mit/für* eine zuverlässige(re) Verbindung zwischen ihrer Form und Bedeutung (bzw. Funktion) auf. Unter Form verstehe ich an dieser Stelle die einzelne Präposition, als Funktion zählt an dieser Stelle der semantische Gehalt der Präposition. Tabelle 6 fasst die Form-Funktionsverbindungen der Präpositionen *mit/für* sowie *an/auf* innerhalb von Objekt-PP in einem Überblick zusammen.

Form: Präposition	Funktion: Semantischer Gehalt (Objekt-PP)	Form-Funktionsverbindung
<i>mit</i> <i>für</i>	semantisiert	prototypisch und zuverlässig
<i>an</i> <i>auf</i>	desemantisiert	nicht prototypisch und weniger zuverlässig

Tabelle 6: Semantisierte und desemantisierte Präpositionen in Objekt-PP: Form-Funktionsverbindungen.

1.2.1.3 Attributive PP

Im Gegensatz zu adverbialen PP oder Objekt-PP stellen attributive PP (PP in attributiver Funktion) keine eigenständigen syntaktischen Satzglieder dar, sondern repräsentieren Gliedteile. Attributive PP können u.a. auf die bereits oben ausgeführten Satzgliedfunktionen von PP – Objekt-PP und Adverbiale – zurückgeführt werden. Basierend auf dem syntaktischen Kontext und der daraus resultierenden Nähe zu Objekt-PP oder adverbialer PP kann innerhalb attributiver PP folglich Präpositionsreaktion vorliegen sowie der Gebrauch einer desemantisierten Präposition. Attributive PP können aber auch autonome und semantisierte Präpositionen enthalten.

Die Beispiele unten illustrieren attributive PP: Bei *am Finale* (13.) handelt es sich um eine attributive PP, die aufgrund der Substantivierung des Verbs *teilnehmen an* Ähnlichkeiten zur entsprechenden Objekt-PP des Verbs *teilnehmen aufweist*, während *auf dem Tisch* (14.) eine attributive PP repräsentiert, die von einer adverbialen PP (*Das Essen steht auf dem Tisch*) abgeleitet ist.

13. *Die Teilnahme am Finale war gesichert.*

14. *Das Essen auf dem Tisch war kalt.*

Es lässt sich hier folglich zwischen objekt- und adverbialähnlichen PP-Attributen differenzieren. Als objektähnliche PP-Attribute gelten solche, die sich auf Objekte zurückführen lassen. Hierzu gehören nicht nur Objekt-PP, sondern auch Kasusobjekte, vgl. die unteren Beispiele für Dativ- (15.) oder für Akkusativobjekte (16.), die in Attributfunktion formal als PP realisiert sind:

15.

Das Vertrauen auf die Polizei.

Der Dank an meinen Lehrer.

16.

Das Verlangen nach Gerechtigkeit.

Der Hass auf meinen Nachbarn.

Eisenberg (2013: 264f.) verweist darauf, dass die Mehrheit der Konstruktionen auf Verben zurückgehen, die parallel zum Kasusobjekt im Dativ oder Akkusativ ebenfalls Objekt-PP regieren können. So kann zum Beispiel das Verb *verlangen* mit einem Akkusativobjekt oder mit einer Objekt-PP gebraucht werden, vgl. die folgenden Beispiele:

Ich verlange meinen Anwalt.

Ich verlange nach meinem Anwalt.

Jedoch ist die parallele Existenz von Kasusobjekten und Objekt-PP, wie sie oben vorliegt, selten gegeben. So leiten sich die attributiven PP *an meinen Lehrer* und *auf meinen Nachbarn* (s. oben, 15./16.) nicht von entsprechenden Objekt-PP ab. Die Verben *danken* und *hassen* regieren jeweils ausschließlich Kasusobjekte, die Substantivierungen dieser Verben nehmen jedoch PP als Attribute, vgl. dafür die folgenden Beispiele:

Ich danke meinem Lehrer.

Ich hasse meinen Nachbarn.

Der Vollständigkeit halber sei an dieser Stelle erwähnt, dass sich attributive PP ebenfalls auf das Subjekt eines Satzes beziehen können. So können PP mit den Präpositionen *von* und *durch* auf Subjekte zurückgeführt werden, siehe dazu die folgenden Beispiele (vgl. Eisenberg 2013: 267):

Der Bau von Mietshäusern durch die Stadt²⁹.

Der Dank von Michael.

Manche attributive PP lassen sich dahingegen nicht auf spezifische Satzgliedfunktionen zurückführen. Dazu gehören attributive PP bei Substantiven ohne verbale Basis, vgl. *das Recht auf Vertagung* (Breindl 1989: 14).

Als letzte Klasse attributiver PP seien noch solche genannt, die Varianten zu Genitivattributen darstellen (vgl. die unteren Beispiele 17.-20.). Duden (2016) nennt hier den possessiven Genitiv (837ff.) (17.), den Genitiv Qualitatis (843f.) (18.), den partitiven Genitiv (844f.) (19.) sowie die partitive Apposition, die eng mit dem partitiven Genitiv verwandt ist (844/993f.) (20.):

²⁹ In diesem Beispiel haben wir es mit zwei attributiven PP zu tun. Die PP *von Mietshäusern* stellt ein objektähnliches PP-Attribut dar, da es auf das Akkusativobjekt des Verbs *bauen* zurückgeht. Die PP *durch die Stadt* ist auf das Subjekt zurückzuführen.

17. *Von meiner Schwester der Freund ist Gärtner.*
18. *Ein Mensch mit gutem Willen kann das lösen.*
19. *Die Hälfte vom Kuchen war schon weggegessen.*
20. *Der Bericht führt eine erstaunliche Anzahl an neuen Projekten auf.*

Attributive PP wie solche in 17. sind vor allem typisch für den mündlichen Sprachgebrauch (vgl. Duden 2016: 840). Die restlichen attributiven PP (18.-20.) sind dahingegen auch im schriftlichen Sprachgebrauch üblich. Die bisherigen Ausführungen zeigen die semantische Vielfalt attributiver PP, die sich z. T. aus ihren verbalen Ableitungen ergibt. Die syntaktische Funktion Attribut ist in der Tat diejenige, in der eine „Präposition mit jeder ihrer Bedeutungen auftreten kann“ (Eisenberg 2013: 263).

Als Gliedteile sind attributive PP in NP eingebettet und einige dieser attributiven PP sind „in der Semantik des Substantivs vorangelegt“ (Duden 2016: 812), hier liegt dann Präpositionsrektion vor. Duden (2016: 812) unterscheidet zwar zwischen attributiven PP, die in der Bedeutung des Substantivs vorangelegt sind und nicht weggelassen werden können, und attributiven PP, die in der Bedeutung des Substantivs nicht vorangelegt sind und weggelassen werden können, macht aber auch gleichzeitig auf die Unschärfe dieser Trennung aufmerksam (vgl. Duden 2016: 156). Eisenberg (2013: 268) lehnt diesbezüglich eine Unterteilung zwischen weglassbaren und nicht weglassbaren attributiven PP grundsätzlich ab (vgl. Eisenberg 2013: 268).

In der vorliegenden Arbeit werden attributive PP nicht näher differenziert. Sie werden jedoch als Gliedteile von den syntaktischen Funktionen Objekt-PP und adverbiale PP abgegrenzt. In den empirischen Untersuchungen dieser Dissertation, in denen eine manuelle Datenaufbereitung stattfindet, wird der Gebrauch attributiver PP aufgrund ihrer semantischen Vielfältigkeit und ihres Status als Gliedteil nicht näher betrachtet.

1.2.2 Zusammenfassung der zentralen Terminologie II

Die folgende Tabelle (Tab. 7) gibt eine Zusammenfassung der zentralen Terminologie dieses Kapitels.

Term		Erläuterung
Syntaktische Funktion der PP	Adverbiale PP	PP in der Funktion eines Adverbiales Präposition: i.d.R. autonom (keine Präpositionsreaktion), i.d.R. semantisiert Beispiele: a) <i>Ich lese ein Buch <u>auf dem Balkon</u>.</i> b) <i>Ich wohne <u>in Mannheim</u>.</i>
	Objekt-PP	PP in der Funktion eines Präpositionalobjekts o. Prädikativs ³⁰ Präposition: regiert (Präpositionsreaktion), desemantisiert o. semantisiert Beispiele: a) <i>Petra kümmert sich <u>um das Auto</u>.</i> b) <i>Hans stimmt <u>für Partei X</u>.</i> c) <i>Sie erklärte den Patienten <u>für gesund</u>.</i> d) <i>Sie ist stolz <u>auf ihre Tochter</u>.</i>
	Attributive PP	PP in der Funktion eines Attributs. Ein Attribut ist ein Gliedteil innerhalb einer Nominalphrase. Präposition: regiert oder autonom; semantisiert oder desemantisiert (je abhängig von verbaler Basis, wenn vorhanden) Beispiele: a) <i>Die Teilnahme <u>am Finale</u></i> b) <i>Der Dank <u>an meinen Lehrer</u></i> c) <i>Das Essen <u>auf dem Tisch</u></i> d) <i>Eine große Anzahl <u>an leistungsfähigen Rechnern</u></i>
Kasushierarchie		Angenommene Hierarchie der Kasus Nominativ – Akkusativ – Dativ – Genitiv und Objekt-PP (Eisenberg 2013: 63). Die Kasushierarchie spiegelt die Markiertheitsverhältnisse der einzelnen Kasusformen wider. Darunter zählen die formale Kodierung der Formen und die Distribution (Frequenz) der durch die Formen realisierten syntaktischen Funktionen. Innerhalb der Kasushierarchie nimmt die Objekt-PP eine Randstellung ein.

Tabelle 7: Präpositionalphrasen aus syntaktischer Perspektive: Zentrale Terminologie.

1.2.3 PP unter valenzgrammatischer Perspektive

Valenz wird im Duden (2016: 779) definiert als Eigenschaft eines Wortes, Ergänzungen zu fordern. Eine *Ergänzung* wird dabei verstanden als „eine Phrase, die in der Semantik eines

³⁰ Unter Objekt-PP werden die syntaktischen Funktionen von PP Präpositionalobjekt und Prädikativ, zusammengefasst. Zu Objekt-PP zählen ebenfalls PP, die von Adjektiven regiert werden.

Wortes vorangelegt ist“ (Duden 2016: 779). Formal betrachtet können Ergänzungen unterschiedlich realisiert werden (vgl. Duden 2016: 779f.). In Bezug auf die Anzahl der Ergänzungen spricht man auch von *Stelligkeit*. Das Gegenstück zur *Ergänzung* bildet „eine Phrase, die ein Wort, eine Phrase oder den Satz modifiziert und kommentiert“ (Duden 2016: 779). Hierfür wird der Begriff *Angabe* gebraucht. Eine Angabe ist demnach nicht in der Bedeutung eines Wortes vorangelegt. Handelt es sich bei dem Wort, das Ergänzungen fordert, um ein Verb, spricht man auch von Verbvalenz.

In diesem Kapitel werden Präpositionen und PP unter valenzgrammatischer Perspektive betrachtet. Im Fokus steht die syntaktische Valenz (1.2.3.1) und die semantische Valenz von PP (1.2.3.2).

1.2.3.1 Syntaktische Valenz: PP als Ergänzungen und Angaben

Unter dem Begriff *syntaktische Valenz* wird hier primär die formale und funktionale Realisierung von Ergänzungen verstanden, also ihre Form und syntaktische Funktion. Nach Eisenberg (2013: 63) wird die syntaktische Verbvalenz „spezifiziert in Hinsicht auf Stelligkeit, formale Restriktionen für die einzelne Stellen (Rektion) und Fakultativität/Obligatorik der Besetzung jeder einzelnen Stelle oder bestimmter Stellenkombination“, d. h. zwischen weglassbaren und nichtweglassbaren Ergänzungen (vgl. Duden 2016: 781). Auch weglassbare Ergänzungen sind semantisch im Verb vorangelegt, ihre Weglassbarkeit führt lediglich dazu, dass sie oberflächlich implizit bzw. „sprachlich unausgedrückt bleiben“ (Duden 2016: 781), jedoch mitverstanden werden. Die weglassbare Ergänzung kann somit entfallen, ohne dass der Satz ungrammatisch wird. Die Unterscheidung zwischen weglassbaren und nichtweglassbaren Ergänzungen ist „allerdings nicht völlig strikt“ (Duden 2016: 781) und oft nicht eindeutig. Für einen Überblick zu verschiedenen Abgrenzungskriterien und zu den verschiedenen Formen der Weglassbarkeit von Ergänzungen s. Storrer (1996, 2003).

Das Thema Valenz wird in verschiedenen Monographien, Sammelbänden, Aufsätzen und Handbüchern ausführlich behandelt (vgl. u. a. Tesnière 1980, José-Domínguez-Vázquez/Eichinger 2015, Storrer 1992, Jacobs 2003, Eichinger/Kubczak/Berens 2011, Willems 2006, Ágel et al. 2003, 2006). In diesem Kapitel beschränke ich mich vorrangig auf die Ausführungen in Duden (2016) und Eisenberg (2013). Im Folgenden wird ausschließlich die Verbvalenz behandelt. Im Fokus steht dabei die Klassifikation von PP als Ergänzungen und Angaben.

PP als Ergänzungen

Unter einer PP-Ergänzung verstehe ich eine PP, die den Status einer Verbergänzung hat. Im folgenden Beispiel liegt solch eine PP-Ergänzung vor:

Er denkt an seine Mutter.

Die PP-Ergänzung *an seine Mutter* wird vom Verb *denken* regiert. Die Form einer PP wird somit vom Verb festgelegt, es liegt Präpositionsrektion mit einer festen Kasusrektion vor. Unter einer festen Kasusrektion verstehe ich hier, dass die Präposition *an* innerhalb der PP-

Ergänzung einen festen Kasus regiert (hier: Akkusativ), auch wenn dieselbe Präposition in lokaler Relation zwei Kasus (Akkusativ und Dativ) regieren kann (vgl. Kap. 1.1.1).

Die PP-Ergänzung repräsentiert auf Ebene der syntaktischen Funktionen eine Objekt-PP – in diesem Beispiel können die Begrifflichkeiten PP-Ergänzung und Objekt-PP also gleichgesetzt werden. Die Ebenen syntaktische Funktion und Verbvalenz sind jedoch voneinander zu trennen, da sie „zwei eigenständige Perspektiven“ darstellen (Duden 2016: 793). Eine PP-Ergänzung repräsentiert nicht immer eine Objekt-PP, sondern kann z. B. auch die syntaktische Funktion eines Adverbials einnehmen, vgl. das folgende Beispiel (entnommen aus Duden 2016: 798):

Otto begibt sich auf den Balkon.

Die PP *auf den Balkon* ist eine PP-Ergänzung, die nicht weglassbar ist (**Otto begibt sich*). Sie referiert auf die Richtung der Handlung, die durch das Verb ausgedrückt wird, und ist somit in der Bedeutung des Verbs vorangelegt. Die Realisierung der Richtung als PP ist dabei eine Möglichkeit, diese Ergänzung formal zu realisieren. Sie ist jedoch formal nicht festgelegt, d. h. u.a., dass keine Präpositionsreaktion vorliegt (vgl. *Otto begibt sich in die Stadt/an seinen Platz*)³¹. Die Klassifikation einer PP als PP-Ergänzung ist damit nicht gleichzusetzen mit der syntaktischen Funktion einer PP. Objekt-PP und adverbiale PP können beiderseits PP-Ergänzungen repräsentieren.

PP als Angaben

Unter einer PP-Angabe verstehe ich eine PP, die den Status einer Angabe hat. Wie oben erwähnt, sind Angaben nicht in der Bedeutung eines Verbs vorangelegt. Im folgenden Beispiel ist *im Keller* eine PP-Angabe:

Er öffnete die Weinflasche im Keller.

Da Angaben nicht in der Bedeutung des Verbs angelegt sind, wird ihr Gebrauch nicht vom Verb gesteuert, sie sind demnach formal nicht festgelegt und es liegt keine Präpositionsreaktion vor (vgl. *Er öffnete die Weinflasche auf der Dachterrasse/Er öffnete die Weinflasche hier*). Bezüglich der beiden letzten Aspekte liegen also ähnliche Verhältnisse vor wie bei adverbialen PP-Ergänzungen (s. o.).

Die Grenze zwischen PP-Ergänzungen und PP-Angaben ist nicht trennscharf, da auch die Probe auf Weglassbarkeit nicht immer ein sicheres Indiz darstellt (vgl. Storrer 2003). Zur Bestimmung von Verbergänzungen und zur Abgrenzung von Ergänzungen und Angaben stellen Valenzwörterbücher wie z. B. das elektronische Valenzwörterbuch deutscher Verben (E-

³¹ Des Weiteren ist die Form einer PP hier nicht obligatorisch. Die Richtung kann formal ebenfalls in Form einer Adverbphrase realisiert werden: *Otto begibt sich dorthin*.

Valbu³²) (Kubczak 2009) eine hilfreiche Ressource dar. E-Valbu bietet zu ausgewählten Verben bzw. Verblesarten u.a. Informationen zur syntaktischen Valenz. Je Lesart sind die einzelnen Verbergänzungen in einem Satzbauplan spezifiziert. PP-Ergänzungen, die die Funktionen Objekt-PP oder adverbiale PP einnehmen, werden dabei gesondert klassifiziert³³. PP, die im Satzbauplan einer Verblesart nicht aufgeführt werden, sind als Angaben einzustufen.

Für die Datenaufbereitung der empirischen Untersuchungen orientiere ich mich vor allem an den Satzbauplanangaben in E-Valbu. Diese dienen ebenfalls dazu, PP-Ergänzungen in der Funktion einer Objekt-PP von PP-Ergänzungen in der Funktion eines Adverbiales abzugrenzen. Im empirischen Teil der vorliegenden Dissertation unterscheide ich nicht zwischen PP-Ergänzungen und PP-Angaben, im Vordergrund stehen die syntaktischen Funktionen von PP, unabhängig von ihrem Valenzstatus.

1.2.3.1 Semantische Valenz: Semantische Rollen von PP

Syntaktische und semantische Valenz repräsentieren verschiedene Dimensionen des Valenzbegriffs (vgl. Primus 2012: 15). Geht es bei der syntaktischen Valenz u.a. um formale Restriktionen (Rektion) und die funktionale Realisierung einzelner Ergänzungen, stehen bei der semantischen Valenz die semantischen Rollen im Zentrum, die z. B. von einem Verb auf Basis seiner Bedeutung vergeben werden (s. auch Engelberg 2003: 53). Eine semantische Rolle repräsentiert die Bedeutungsseite einer Ergänzung (Eisenberg 2013: 70, vgl. auch Duden 2016: 397f.). So vergibt das Verb *unterschreiben* zwei semantische Rollen: Eine Rolle des Handlungsträgers (Agens) und die Rolle des Gegenstands, der direkt von der Handlung betroffen ist, ohne eine Kontrolle auf diese ausüben zu können (Patiens) (vgl. Duden 2016: 398). Somit ist hier das Verb Valenzträger (Primus 2012: 2) und vergibt als solcher zwei semantische Rollen, Agens und Patiens: *Die Anwältin (Agens) unterschreibt den Brief (Patiens)*. Semantische Rollen können formal unterschiedlich realisiert sein. Die semantische Rolle einer Entität lässt sich nur im syntaktischen Kontext ermitteln. In der Regel werden semantische Rollen durch Verben vergeben, aber auch Adjektive können semantische Rollen vergeben und damit Valenzträger sein (vgl. Duden 2016: 792). Semantische Rollen werden auch semantische Argumente genannt (Primus 2012: 2). Die Argumentstruktur eines Verbs bedingt somit formale und semantische Beschränkungen von Verbergänzungen (Engelberg et al. 2011: 71).

Bei Entitäten, die keine semantischen Argumente darstellen, handelt es sich um Modifikatoren (vgl. Primus 2012: 2). Die Unterteilung in semantische Argumente und Modifikatoren entspricht auf syntaktischer Ebene der Unterteilung in Ergänzungen und Angaben: Semantische Argumente sind Ergänzungen³⁴, Modifikatoren sind Angaben. Tabelle 8 illustriert den Argumentstatus von PP (nach Primus 2012: 4). Unter einem PP-Argument verstehe ich PP mit Status eines semantischen Arguments, ein PP-Modifikator ist folglich eine PP mit Status eines Modifikators.

³² Seit Dezember 2018 verfügbar unter <https://grammis.ids-mannheim.de/verbvalenz>.

³³ Die entsprechenden Klassen lauten Präpositionsergänzung (bzw. Präpositivkomplement) und Prädikativergänzung (bzw. Prädikativkomplement) sowie Adverbativergänzung (bzw. Adverbialkomplement), vgl. E-Valbu.

³⁴ Primus (2012: 4) nennt hier auch den Begriff *syntaktisches Argument*.

PP-Status		Beispiel und Erläuterung
1	PP-Ergänzung (Objekt-PP)	<i>Oma achtet <u>auf ihre Gesundheit</u></i> Semantische Rollenzuweisung PP-Argument
2	PP-Ergänzung (Adverbiale)	<i>Opa wohnt <u>auf dem Dachboden</u></i> Semantische Rollenzuweisung Raumergänzung PP-Argument
3	PP-Angabe (Adverbiale)	<i>Max spielt <u>auf dem Dachboden</u></i> Keine semantische Rollenzuweisung Raumangabe PP-Modifikator

Tabelle 8: Argumentstatus von PP (nach Primus 2012: 4).

Tabelle 8 zeigt, dass eine adverbiale PP-Ergänzung ein semantisches Argument darstellen und damit eine durch den Valenzträger zugewiesene semantische Rolle füllen kann (Beispiel 2). Für adverbiale PP-Angaben trifft dies jedoch nicht zu (Beispiel 3). In beiden Beispielen (2 und 3) liegt jedoch dieselbe syntaktische Funktion sowie dieselbe semantische Rolle vor. Primus (2012: 62) verweist hier auf die Lokalrolle, die zu den Adverbialrollen gezählt wird. Als Adverbialrollen gelten semantische Argumente in der syntaktischen Funktion eines Adverbiales. Eine Lokalrolle entspricht hier also einem Lokaladverbiale.

Nach Primus (2012) lassen sich sowohl für Ergänzungen als auch Angaben semantische Rollen beschreiben, auch wenn für Angaben angenommen wird, dass diese gegenüber den semantischen Rollen von Ergänzungen als hierarchisch niedriger einzustufen sind (vgl. Breindl 1989: 53). In Bezug auf PP bedeutet dies z. B., dass adverbiale PP unabhängig von ihrem Status als PP-Ergänzung oder PP-Angabe mit semantischen Rollen beschrieben werden können (s. Tab. 8). Da PP im Fokus meiner Dissertation stehen, gehe ich im Folgenden nur auf semantische Rollen von PP ein. Hierbei beschränke ich mich vorrangig auf die semantischen Rollen, die Objekt-PP³⁵ und adverbiale PP kodieren können.

Eine ausführliche Gegenüberstellung semantischer Rollen von Objekt-PP und adverbialer PP findet man in der Arbeit von Breindl (1989). Breindl (1989) verzichtet auf eine vorausgehende systematische Aufzählung semantischer Rollen, stellt aber in Hinsicht auf die syntaktischen Funktionen Objekt-PP und adverbiale PP fest, dass diese zumindest die in Tabelle 9 angeführten semantische Rollen gleichermaßen kodieren können. Die in der Tabelle enthaltenen Beispiele stammen aus Breindl (1989: 54f.), die PP ist jeweils unterstrichen.

³⁵ Unter Objekt-PP fasse ich für die Dissertation Präpositionalobjekte und PP in prädikativer Funktion zusammen. Semantische Rollen prädikativer PP finden in Primus (2012) sowie in Eisenberg (2013) jedoch keine explizite Erwähnung.

Semantische Rolle	Beispiel adverbiale PP (A) und Objekt-PP (O)	
Ausgangspunkt ³⁶	A	Werner kommt gerade <u>vom Bahnhof</u> .
	O	Werner stammt <u>von einem alten Adelsgeschlecht</u> ab.
Ziel	A	Werner schaut/blickt <u>auf den Boden</u> .
	O	Werner schaut/achtet nicht <u>auf eine vitaminreiche Ernährung</u> .
Ort	A	Werner steht nicht mehr ganz fest <u>auf dem Boden</u> .
	O	Alle stehen <u>auf Werner</u> .
Benefaktiv	A	Werner arbeitet <u>für/gegen seine eigenen Kollegen</u> .
	O	Werner entscheidet sich <u>für/gegen den Hauptkandidaten</u> .
Qualität	A	<u>Aus der Wurzel</u> hat der Sepp dann einen Gartenzwerg geschnitzt.
	O	Wasser besteht <u>aus Sauerstoff- und Wasserstoffatomen</u> .
Grund	A	Werner leidet sehr <u>wegen seiner Leberzirrhose</u> .
	O	Werner leidet <u>an Leberzirrhose</u> .
Begleiter	A	Werner singt <u>mit Helga</u> ein Liedchen.
	O	Werner bequatscht alles <u>mit Helga</u> .
Instrument	A	Werner belädt den Wagen <u>mit der Heugabel</u> .
	O	Werner begründet seinen Kreditantrag <u>mit chronischem Finanzmangel</u> .

Tabelle 9: Ähnliche semantische Rollen bei Objekt-PP und adverbialen PP: Beispiele (nach Breindl 1989: 54f.).

Die Beispiele in Tabelle 9 zeigen: In Bezug auf semantische Rollen stehen sich eine Vielzahl von Objekt-PP und adverbiale PP semantisch „nahe“ (Breindl 1989: 54). Auch im Duden (2016: 401) findet man Hinweise auf die Ähnlichkeit semantischer Rollen von PP-Ergänzungen. So stellt Duden (2016: 401) fest, dass semantische Rollen von PP-Ergänzungen „einen mehr oder weniger durchsichtigen Zusammenhang mit prototypischen Anwendungen der jeweiligen Präpositionen aufweisen“. Genannt werden in diesem Zusammenhang u.a. die semantischen Rollen *Werkzeug*, *Ziel* oder *Quelle*, s. die unteren Beispiele in entsprechender Reihenfolge. Dabei differenziert Duden (2016: 401) hier nicht nach den syntaktischen Funktionen der PP, in den Beispielen werden sowohl adverbiale als auch Objekt-PP aufgeführt³⁷.

jmdn. mit etw. töten

jmdn. an etwas erinnern

jmdn./etw. von jmdn./etw. trennen

Neben einer Ähnlichkeit zu semantischen Rollen adverbialer PP finden sich bei Objekt-PP ebenso Ähnlichkeiten zu den semantischen Rollen reiner Kasusobjekte, beide Konstruktionen

³⁶ In Bezug auf die semantische Rolle Ausgangspunkt bzw. Source erwähnt Breindl (1989: 62), dass hier vorrangig die Präpositionen *von* und *aus* auftreten.

³⁷ Nach E-Valbu handelt es sich beim ersten Beispiel (*jmdn. mit etwas töten*) um eine adverbiale PP-Ergänzung, beim zweiten und dritten Beispiel (*jmdn. an etwas erinnern* bzw. *jmdn./etw. von jmdm./etw. trennen*) um ein Präpositivkomplement, folglich um eine Objekt-PP.

können sogar alternieren (s. auch Proost 2015). Breindl (1989: 98) nennt hier u. a. folgendes Beispiel:

Max schreibt Onkel Fritz/an Onkel Fritz einen Brief

Das Beispiel veranschaulicht, dass die semantische Rolle *Rezipient/-in* sowohl durch Objekt-PP auch durch reine Kasusobjekte kodiert werden kann. Das Beispiel dient ebenfalls als Illustration der grundlegenden Annahme zu Objekt-PP, dass eine PP nur dann ein PP-Objekt ist, „wenn dieselbe semantische Rolle als nominales Objekt kodiert sein kann“ (Eisenberg 2013: 302). Diese Bedingung ist jedoch nicht immer eindeutig überprüfbar und es wird gar angezweifelt, ob Objekt-PP und reine Kasusobjekte tatsächlich identische semantischen Rollen kodieren (Eisenberg 2013: 303). Nach Breindl (1989: 57) sind semantische Kodierungen von Objekt-PP im Vergleich zu reinen Kasusobjekten „nicht völlig bedeutungsgleich“. Breindl (1989: 56ff.) nimmt für semantische Rollen eher ein Kontinuum an, auf dem sich Objekt-PP und reine Kasusobjekte als unterschiedliche Ausprägungen einzelner semantischer Rollen einordnen lassen. Eine Ähnlichkeit der semantischen Rollen von Objekt-PP und reinen Kasusobjekten wird somit nicht abgelehnt. Im Deutschen gibt es viele Fälle, in denen z. B. PP- und Dativobjekte alternieren können, auch wenn gewisse semantische Einschränkungen zu beachten sind (s. dazu z. B. Primus 2012: 65ff.).

Eine Alternation von NP und PP ist ebenfalls bei der semantischen Rolle *Benefaktiv* oder *Quelle* möglich (Primus 2012: 67ff.), s. die folgenden Beispiele:

Oma strickt ihrem Enkel einen Pulli/einen Pulli für ihren Enkel. (Benefaktiv)

Ein Einbrecher stahl dem Opa Geld/vom Opa Geld. (Quelle)

In Bezug auf Dativ-NPs verweisen Helbig/Buscha (2001b: 264) darauf, dass der Dativus commodi (ein Untertyp des freien Dativs³⁸) „eine starke Ähnlichkeit zum Objekt“ zeigt, jedoch (anders als das Dativobjekt) durch eine PP mit der Präposition *für* ersetzt werden kann (Helbig/Buscha 2001b: 464, s. vorangehendes Beispiel). Das bedeutet, dass es im Deutschen sprachliche Kontexte gibt, in denen eine Dativ-NP durch eine *für*-PP ausgetauscht werden kann, ohne dass dies einen deutlichen semantischen Unterschied erzeugt.

Bezüglich der semantische Rolle Agens erwähnt Breindl (1989: 62), dass komitative *mit*-PP, die belebte Denotate spezifizieren, unabhängig von ihrem Status als PP-Ergänzung oder PP-Angabe der semantischen Rolle des Agens „sehr nahe“ stehen, s. dazu die folgenden Beispiele aus Breindl (1989: 61):

Innstetten duelliert sich mit Crampas.

Max kommt mit Moritz zum Abendessen.

³⁸ Helbig/Buscha (2001b: 464) sehen im freien Dativ eine „starke Ähnlichkeit“ zum Dativobjekt, zählen freie Dative jedoch zu sekundären Satzgliedern, d. h. zu solchen Satzgliedern, die keine Ergänzungen zum Verb darstellen. Auch Duden (2016: 829) versteht unter einem freien Dativ eine Dativ-NP, die keine Ergänzung zum Verb darstellt, zählt freie Dative jedoch zu den Dativobjekten.

Hinweise auf typische semantische Rollen adverbialer PP findet man in Eisenberg (2013: 70f.) und Primus (2012: 62ff.). In seiner Auflistung der semantischen Rollen gibt Eisenberg (2013: 70f.) unter den Rollen *Lokativ*, *Direktiv* und *Instrumental* lediglich Beispiele mit adverbialen PP an. Implizit wird so deutlich, dass zumindest adverbiale PP scheinbar prototypisch diese semantischen Rollen kodieren. Auch bei Primus (2012: 62ff.) findet man adverbiale PP vorrangig unter der Abhandlung von Lokal- und Temporalrollen sowie den Rollen Instrument, Komitativ, die ihrerseits sowohl durch adverbiale PP-Ergänzungen als auch durch adverbiale PP-Angaben versprachlicht werden können. Beispiele für die genannten Rollen nach Primus (2012: 63-78) sind im Folgenden aufgeführt:

Lokalrolle: *Max wohnt in einem Dorf.*

Temporalrolle: *Die Vorlesung ist am Montag.*

Instrument: *Der Einbrecher öffnete die Tür rücksichtslos mit einem Brecheisen.*

Komitativ: *Der Einbrecher telefoniert mit seinem Komplizen.*

Nach Primus (2012) liegen die semantischen Rollen *Instrument* und *Komitativ* semantisch nah bei einander und verweisen auf „Mitbeteiligung“ (Primus 2012: 73). Temporalrollen werden nach Primus (2012: 79) „analog zu Lokalrollen konzeptualisiert und syntaktisch realisiert“. Daraus kann geschlussfolgert werden, dass auch hier die Kodierung entsprechender semantischer Rollen durch adverbiale PP typisch ist.

Zusammenfassend lässt sich hier festhalten, dass adverbiale und Objekt-PP ähnliche semantische Rollen kodieren können bzw. Heterogenität semantischer Rollen bei beiden syntaktischen Funktionen vorliegt. Für einen Großteil der Objekt-PP zeigt sich in Bezug auf semantische Rollen nach Breindl (1989: 66) eine „deutliche Nähe zum Adverbialsystem“, und damit zu adverbialen PP.

1.2.4 Zusammenfassung der zentralen Terminologie III

Tabelle 10 fasst die zentralen Termini zu PP aus valenzgrammatischer Perspektive zusammen.

Term	Erläuterung
Valenz	Eigenschaft eines Wortes, eine Ergänzung/Ergänzungen zu fordern. Verbvalenz = Verb fordert eine Ergänzung/Ergänzungen.
Syntaktische (Verb-)Valenz	Formale und funktionale Realisierung von (Verb-)Ergänzungen.
Semantische (Verb-)Valenz	Semantische Perspektive auf (Verb-)Ergänzungen.
Rektion	Formale Restriktionen für einzelne Ergänzungen Beispiele: Das Verb <i>helfen</i> regiert eine Ergänzung im Dativ (<i>Ich helfe dir</i>) Das Verb <i>kümmern</i> regiert eine PP-Ergänzung mit der Präposition <i>um</i> (<i>Er kümmert sich um die Kinder</i>).
Ergänzung	„eine Phrase, die in der Semantik eines Wortes vorangelegt ist“ (Duden 2016: 779), d. h. von diesem Wort (z. B. Verb) regiert wird.
Angabe	„eine Phrase, die ein Wort, eine Phrase oder den Satz modifiziert und kommentiert“ (Duden 2016: 779), d. h. nicht von einem Wort (z. B. Verb) regiert wird.
PP-Ergänzung	PP mit Status einer Ergänzung
PP-Angabe	PP mit Status einer Angabe
Semantische Rolle	Die Bedeutungsseite einer Ergänzung. Beispiel: Semantische Rolle <i>Agens</i> verweist auf die Rolle eines Handlungsträgers.
semantisches Argument	Entität in einem Satz, der eine semantische Rolle vom Valenzträger zugewiesen wird.
Modifikator	Entität in einem Satz, der keine semantische Rolle vom Valenzträger zugewiesen wird.
PP-Argument	PP mit Status eines semantischen Arguments
PP-Modifikator	PP mit Status eines Modifikators
E-Valbu	Das elektronische Valenzwörterbuch deutscher Verben. Seit Dez. 2018 verfügbar unter https://grammis.ids-mannheim.de/verbvalenz (Stand: März 2019). E-Valbu enthält Informationen zu Ergänzungen ausgewählter Verben bzw. Verblesarten.

Tabelle 10: Präpositionalphrasen unter valenzgrammatischer Perspektive: Zentrale Terminologie.

In der vorliegenden Arbeit liegt der primäre Fokus auf (regierten) Präpositionen innerhalb von PP der syntaktischen Funktionen Objekt-PP und Adverbiale, unabhängig von ihrem Valenzstatus. Aus diesem Grund erfolgt in der Arbeit keine Trennung zwischen PP-Ergänzungen und PP-Angaben. Das bedeutet im Besonderen, dass adverbiale PP unabhängig von ihrem Status als PP-Ergänzung oder PP-Angabe als *eine* Gruppe aufgefasst und Objekt-PP

gegenübergestellt werden³⁹. Weiterhin werde ich innerhalb der syntaktischen Funktionen von PP nicht nach einzelnen semantischen Rollen differenzieren, da es, wie oben ausgeführt, zwischen Objekt-PP und adverbialer PP viele semantische Überlappungen geben kann.

1.2.5 Andere Verwendungsweisen von PP

In diesem Kapitel möchte ich der Vollständigkeit halber auf weitere Verwendungsweisen von Präpositionen bzw. PP eingehen. Hierzu zähle ich Funktionsverbgefüge, Komparationsformen von Adjektiven und die Verb-Verlaufsform sowie formelhafte Sequenzen. Im Folgenden gehe ich kurz auf diese Verwendungsweisen ein. Die aufgeführten Verwendungsweisen wurden in den empirischen Fallstudien, in denen eine manuelle Datenaufbereitung erfolgte, von der weiteren Auswertung ausgeschlossen.

1.2.5.1 Funktionsverbgefüge

Unter Funktionsverbgefügen (FVG) verstehe ich in Anlehnung an Duden (2016: 422ff.) mehrteilige Prädikate, die aus einem Funktionsverb und einem deverbalen Substantiv im Akkusativ oder einer PP mit deverbalem Substantiv gebildet werden, s. die unteren Beispiele. Im Folgenden möchte ich mich ausschließlich auf FVG mit PP (s. Beispiel 22) beschränken.

21. *Anwendung finden, (jmdm.) ein Versprechen geben*

22. *zum Abschluss bringen, in Anspruch nehmen*

Als Funktionsverben zählen z. B. *bringen* oder *nehmen* (vgl. Duden 2016: 423). In Verbindungen wie den obigen Beispielen dienen die Funktionsverben zum Ausdruck reiner grammatischer bzw. morphosyntaktischer Merkmale wie Tempus, Modus, Numerus. Den lexikalischen Kern eines FVG bilden jeweils die präpositional angeschlossenen Substantive.

Verbindungen aus Funktionsverb und einem präpositional angeschlossenen Substantiv enthalten Präpositionen, die nicht ausgetauscht werden können. Das Inventar an Präpositionen, die als Bestandteil von FVG gebraucht werden, ist sehr beschränkt, es treten vorwiegend die Präpositionen *in (im)*, *zu (zum, zur)* und *unter* auf⁴⁰ (Duden 2016; 429), s. die folgenden Beispiele:

in Anspruch nehmen

zur Anklage kommen

unter Beobachtung stehen

FVG mit präpositional angeschlossenen Substantiven können als Äquivalente zu Prädikatsausdrücken angesehen werden: Das Funktionsverb bildet mit seinem lexikalischen Kern, der PP, eine Einheit, die dem eines einfachen Vollverbs entspricht (vgl. Duden 2016: 425). Somit stellt das FVG als Ganzes ein mehrteiliges Prädikat dar.

³⁹ Zur Abgrenzung von PP als adverbiale Ergänzungen und adverbiale Angaben vgl. u. a. Breindl (1989, 2006) und GDS (1997: 1368ff./2114ff.).

⁴⁰ Helbig/Buscha (2001b: 72) nennen auch ein FVG mit *auf*: *auf einen/den Gedanken bringen*. FVG mit den Präpositionen *an*, *mit* oder *für* führen sie nicht auf (H/B 2001: 83).

1.2.5.2 Komparationsform von Adjektiven und Verb-Verlaufsform

PP können Bestandteile von Komparationsformen eines Adjektivs sein. Viele Adjektive können Komparationsformen bilden (Duden 2016: 340f.). Zu den Komparationsformen zählen der *Positiv*, der *Komparativ* und der *Superlativ*, s. das folgende Beispiel:

Positiv: *heiß*

Komparativ: *heißer*

Superlativ: *am heißesten*

Der Superlativ enthält die Verschmelzung *am*. Duden (2016: 375) beschreibt den Superlativ (*am heißesten*) als „eine feste Verbindung mit *am*“. Die Form *am*, die formal der Verschmelzung aus Präposition (*an*) und definitivem Artikel (*dem*) entspricht, wird hier nicht als solche gewertet, vielmehr wird *am* klassifiziert als „eine[-] nicht mehr auflösbare[-] Partikel“ (Duden 2016: 298).

Eine ähnliche Klassifikation erfährt im Duden (2016: 298) ebenfalls die Verlaufsform eines Verbs, z. B. *Die Kinder sind schon am Schlafen*. Verb-Verlaufsformen werden im Deutschen mit dem Verb *sein* und einem Infinitiv mit *am* oder *beim* gebildet (s. Beispiele unten) und sind mit der englischen Progressivform (z. B. *be dancing*) vergleichbar (Duden 2016: 435):

Ich bin am Überlegen

Ich bin beim Aufräumen

Die Verb-Verlaufsform ist nach Duden (2016: 435) typischerweise in der gesprochenen Sprache „weiter verbreitet als in der Standardschriftsprache“.

Beide Konstruktionen (Superlativ und Verlaufsform) enthalten nach Duden (2016: 298) die Partikel *am*, die nicht weiter auflösbar ist. Mit dieser Klassifikation wird der Form *am* in beiden oben genannten Konstruktionen der Status einer Verschmelzung aus Präposition und Artikel indirekt aberkannt. Eisenberg (2013: 191) sieht in diesen Konstruktionen dahingegen Verschmelzungen aus Präposition und Artikel mit einem hohen Grammatikalisierungsgrad.

1.2.5.3 Formelhafte Sequenzen

Unter einer formelhaften Sequenz verstehe ich hier in Anlehnung an Wray (2002: 9) “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (Wray 2002: 9, zitiert in Wisniewski et al. 2014: 26)⁴¹. In Bezug auf PP verstehe ich die Definition als Verweis auf PP als Bestandteil vorgefertigter Wortsequenzen unterschiedlichen Typs. Zu PP in formelhaften Sequenzen zähle ich vorrangig Idiome, bei denen sich die Bedeutung der Gesamt-Wortsequenz nicht anhand der Bedeutungen der sequentiellen Einheiten ableiten lässt. Die

⁴¹ Diese Definition findet sich im Annotationsschema des Lernerkorpus MERLIN, das als Datengrundlage der empirischen Untersuchungen meiner Dissertation dient.

Bedeutung von Idiomen ist lexikalisiert (vgl. Burger 2015: 14f., 26f.). So trägt das Idiom *jndn. hinters Licht führen* die Bedeutung *jemanden täuschen* (Duden-Online⁴²). Weitere Beispiele für Idiome sind⁴³:

jndn. auf die Palme bringen
etw. auf die lange Bank schieben

Weiterhin zähle ich hier Beispiele wie die folgenden sowie deren mögliche Abwandlungen⁴⁴ zu den formelhaften Sequenzen:

23.

Mit freundlichen Grüßen
Auf Wiedersehen
Danke im Voraus
*(Vielen) Dank für (deinen Brief)*⁴⁵

24.

Tag für Tag
Schulter an Schulter

Burger (2015: 45) klassifiziert die Gruß- und Abschiedsformeln (23.), „die auf ritualisierte Art den Beginn und den Abschluss [...] markieren“ als Routineformeln (im engeren Sinne). Im Annotationsschema des Lernerkorpus MERLIN fallen derartige Sequenzen unter den Begriff kommunikativer Phraseologismus (Wisniewski et al. 2014: 17). Beispiele unter 24. werden auch Paar- bzw. Zwillingsformeln genannt (Burger 2015: 55).

⁴² Siehe <https://www.duden.de/rechtschreibung/Licht> (März 2019).

⁴³ Die Bedeutungen der Idiome sind in entsprechender Reihenfolge: 1. jemanden aufbringen, wütend machen, erzürnen (<https://www.duden.de/rechtschreibung/Palme>), 2. etwas Unangenehmes aufschieben, hinauszögern (https://www.duden.de/rechtschreibung/Bank_Sitzgelegenheit) (beide: Mai 2019).

⁴⁴ Eine Abwandlung der formelhaften Sequenz *Mit freundlichen Grüßen* wäre z. B. *Mit freundlichem Gruße*, *Mit Gruß* oder *mit schönem Gruß*.

⁴⁵ Die Klammern deuten an, dass diese Elemente austausch- oder weglassbar sind, z. B. *Danke für deinen Brief/Vielen Dank für Ihre Hilfe* etc.

1.2.6 Zusammenfassung der zentralen Terminologie IV

Tabelle 11 fasst die zentrale Terminologie des Kapitels zu anderen Verwendungsweisen von PP zusammen.

Term	Erläuterung
FVG	Funktionsverbgefüge mit einem präpositional angeschlossenen Substantiv Beispiel: <i>zum Abschluss bringen, in Anspruch nehmen</i> FVG mit präpositional angeschlossenen Substantiven werden als Äquivalente zu Prädikatsausdrücken angesehen.
Superlativ	Eine Komparativform bei (einigen) Adjektiven Beispiel: <i>am heißesten</i> Die Verschmelzung <i>am</i> entspricht „einer nicht mehr auflösbare[n] Partikel“ (Duden 2016: 298).
Verlaufsform	Eine Konstruktion mit dem Verb <i>sein</i> und einem Infinitiv mit <i>am</i> oder <i>beim</i> . Beispiel: <i>Die Kinder sind schon am Schlafen</i> Die Bedeutung der Konstruktion ist vergleichbar mit der englischen Progressivform (z. B. <i>be dancing</i>). Die Verschmelzung <i>am</i> entspricht „einer nicht mehr auflösbare[n] Partikel“ (Duden 2016: 298).
Formelhafte Sequenz	“a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (Wray 2002: 9, zitiert in Wisniewski et al. 2014: 26). Zu den formelhaften Sequenzen, die PP enthalten, zähle ich vor allem Idiome (1), Routineformeln (2) sowie Zwillingsformeln (3) (Burger 2015: 26ff., 45, 55). Beispiele: 1) <i>jndn. auf die Palme bringen</i> 2) <i>Mit freundlichen Grüßen/Auf Wiedersehen</i> 3) <i>Tag für Tag/Schulter an Schulter</i>

Tabelle 11: Andere Verwendungsweisen von Präpositionalphrasen: Zentrale Terminologie.

1.3 Präpositionen und Präpositionalphrasen aus kontrastiver Perspektive

Nicht jede Sprache verfügt über Präpositionen und PP wie das Deutsche. Aus sprachkontrastiver Perspektive möchte ich an dieser Stelle unterscheiden zwischen Sprachen, die, ähnlich zum Deutschen, über Präpositionen bzw. PP verfügen und solchen, die über keine Präpositionen bzw. PP verfügen. Ich differenziere somit ausgewählte Sprachen, die der deutschen Sprache typologisch ähnlich sind bzw. sich typologisch vom Deutschen unterscheiden. Unter dem Begriff *typologisch verschieden* (TV) verstehe ich an dieser Stelle

Sprachen, die nicht über prototypische, lexikalische Präpositionen verfügen. Unter *typologisch ähnlich* (TA) verstehe ich dahingegen solche Sprachen, die über eigenständige, lexikalische Präpositionen verfügen und damit in dieser Hinsicht der deutschen Sprache ähnlich sind.

Als typologisch verschieden (TV) fasse ich exemplarisch die Sprachen Türkisch und Ungarisch auf, als typologisch ähnlich (TA) die Sprachen Englisch und Italienisch. Auf beide Gruppen (TV und TA) gehe ich im Folgenden ein, da diese Sprachen für eine Fallstudie dieser Dissertation zentral sind. Ich beschränke mich dabei auf die zentralen Unterschiede im Hinblick auf Präpositionen und PP bzw. die sprachlichen Äquivalente dieser⁴⁶.

In den Sprachen Türkisch und Ungarisch (TV) werden Relationen, die im Deutschen durch Präpositionen und PP ausgedrückt werden, vorrangig mithilfe von Kasussuffixen realisiert (vgl. Schroeder/Şişmek 2010: 721 und Brdar-Szabó 2010b: 733). Im Türkischen stehen hierfür die Kasus Dativ⁴⁷, Ablativ und Lokativ zur Verfügung (vgl. z. B. Kalkavan-Aydin 2017: 376). In Fällen, in denen das Kasussuffix nicht hinreichend die auszudrückenden Relationen spezifiziert, werden Postpositionen⁴⁸ verwendet, die dazu dienen, die ‚allgemeine‘ Bedeutung der Kasussuffixe zu spezifizieren und zu disambiguieren (vgl. Turgay 2010c: 317, Hoffmann 2016: 405). Zusätzlich zählen im Türkischen sogenannte „Ortsnomina“ (Kalkavan-Aydin 2017: 376f.) bzw. „Raumnomen“ (Hoffmann 2016: 405) wie z. B. *ön* (Vorderseite) oder *arka* (Hinterseite) zu den Ausdrücken, die präpositionalen Äquivalenten im Deutschen nahekommen. Das Türkische verfügt nicht über Äquivalente verbregierter Präpositionen (Hoffmann 2016: 404).

Im Ungarischen liegen dem Türkischen ähnliche Möglichkeiten zur Realisierung von Relationen vor, die im Deutschen durch Präpositionen bzw. PP ausgedrückt werden. Auch hier können Kasussuffixe, Postpositionen und Kombination aus beiden eingesetzt werden (vgl. Csillag/Károlyi/Nagy 1983: 230f., Fekete 2008: 173, Forgács 2004: 235f.), wobei Kasussuffixe⁴⁹ den Regelfall darstellen (Forgács 2004: 23). Die Bedeutung der Suffixe ist dabei nicht autonom und hängt vom jeweiligen Kontext ab (Csillag/Károlyi/Nagy 1983: 231, Forgács 2004: 364). Weder im Türkischen noch im Ungarischen gibt es Präpositionen als eigenständige Wortartenkategorie (vgl. Forgács 2004: 69f.). Beispiele aus dem Türkischen und Ungarischen sowie ihre deutschen Übersetzungen sind in Tabelle 12 aufgeführt. Die sprachlichen Mittel der jeweiligen Sprache sind in den Beispielen jeweils unterstrichen.

⁴⁶ Für einen Überblick zum kontrastiven Sprachvergleich Englisch-Deutsch s. Hall (2010), zu Italienisch-Deutsch s. Costa (2010), zu Türkisch-Deutsch s. Schroeder/Şişmek (2010) und zu Ungarisch-Deutsch s. Brdar-Szabó (2010b).

⁴⁷ Interessant ist, dass der Dativ im Türkischen u.a. eine Richtung bzw. die Ausrichtung auf ein Ziel repräsentiert, während der Kasus im Deutschen zum Ausdruck einer statischen Lage verwendet wird (Hoffmann 2016: 404).

⁴⁸ Zur Unterscheidung von primären und sekundären Postpositionen siehe Turgay (2010c: 319ff.). Sie versteht Verbindungen aus Kasussuffix und Ortsnomen als sekundäre Postpositionen. Nach Schroeder/Şişmek (2010: 721) ist im Türkischen jedoch die „Zahl der echten Postpositionen gering“.

⁴⁹ Nach Forgács (2004: 145) werden diejenigen Kasus, die in der Regel Adverbiale (mit einem semantischen Gehalt) ausdrücken, als semantische Kasus klassifiziert während die Kasusformen zur Markierung „grundlegende[r] syntaktische[r] Positionen im Satz wie Subjekt, Akkusativ- bzw. Dativobjekt“ als syntaktische Kasus aufgefasst werden.

Sprache (TV)	Sprachliche Mittel	Beispiel	Übersetzung
Türkisch	Kasussuffixe	<i>Istasyona_ gidiyoruz.</i>	<i>Wir gehen zum Bahnhof.</i>
	Postpositionen	<i>Ali kardesi için çalışıyor.</i>	<i>Ali arbeitet für seinen Bruder.</i>
Ungarisch	Kasussuffixe	<i>A zsebembe teszem a tollat.</i>	<i>Ich stecke den Kugelschreiber in meine Tasche.</i>
	Postpositionen	<i>A lámpát az asztal mellett áll.</i>	<i>Die Lampe steht neben dem Tisch.</i>

Tabelle 12: PP-Äquivalente im Türkischen und Ungarischen: Beispiele⁵⁰.

Wie man Tabelle 12 entnehmen kann, stehen einer prototypisch deutschen Präpositionen bzw. PP sowohl im Türkischen als auch im Ungarischen prototypisch Kasussuffixe und ggf. Postpositionen gegenüber. Ein formgleiches Äquivalent zu prototypischen deutschen Präpositionen und PP gibt es nicht.

Die Sprachen Englisch und Italienisch (TA) verfügen über Präpositionen und PP, die Gemeinsamkeiten mit dem Gebrauch deutscher Präpositionen und PP aufweisen. Das Englische verfügt über Präpositionen, die als prototypisch zum Ausdruck räumlicher Relationen dienen, ein obligatorisches Bezugselement fordern und in der Form einer PP verschiedene syntaktische Funktionen einnehmen können (vgl. Huddleston/Pullum 2006: 603). Die Präpositionen können zum einen eine lexikalische Bedeutung haben (lexikalische Präpositionen) (z. B. *behind*), wobei diese als konkret oder abstrakt klassifiziert werden kann (2011: 74). Zum anderen können Präpositionen auch vom Verb regiert werden und damit grammatikalisiert sein (grammatische Präpositionen) (Huddleston/Pullum 2006: 601). Im Gegensatz zum Deutschen weisen Präpositionen im Englischen keine Kasusreaktion auf.

Die Italienische Sprache verfügt ebenfalls über Präpositionen, die in PP je nach ihrem syntaktischen Vorkommen als lexikalisch (z. B. *in, con*) oder grammatisch (z. B. *di, a*) klassifiziert werden können (Schwarze 2011: 292). Präpositionen und PP können im Italienischen von Verben regiert werden (grammatische Präpositionen⁵¹), drücken jedoch größtenteils semantisch beschreibbare Relationen aus, z. B. Angaben zu Ort, Richtung oder Zeitpunkt (lexikalische Präpositionen⁵²) (Schwarze 2011: 292, 302). Wie im Englischen können PP im Italienischen unterschiedliche syntaktische Funktionen realisieren (Schwarze 2011: 308f.). Beispiele für Präpositionen bzw. PP im Englischen und Italienischen sowie ihre deutschen Übersetzungen sind in Tabelle 13 angegeben. Die sprachlichen Mittel sind in den Beispielen jeweils unterstrichen.

⁵⁰ Die Beispiele (inkl. ihrer Übersetzungen) sind entnommen aus Turgay (2010c: 314/317) und Forgács (2004: 154f./235f.).

⁵¹ Da das Italienische über keine Kasusflexion verfügt, haben einige grammatische Präpositionen die Funktion, die unterschiedlichen Kasus morphematisch zu markieren (Rindler Schjerve 1981: 127).

⁵² Schwarze (2011: 302) erwähnt jedoch, dass lexikalische Präpositionen (z. B. *con*) auch grammatisch verwendet werden können, wenn sie von Verben regiert sind, d. h. wenn Präpositionsreaktion vorliegt (*scusarsi con qu.*, dt. *sich bei jmdm. entschuldigen*).

Sprache (TA)	Sprachliche Mittel	Beispiel	Übersetzung
Englisch	Lexikalische Präposition	<i>Jill is <u>behind</u> her desk</i>	<i>Jill ist hinter ihrem Schreibtisch.</i>
	Grammatische Präposition	<i>It depends <u>on</u> the circumstances.</i>	<i>Es hängt von den Umständen ab.</i>
Italienisch	Lexikalische Präposition	<i>Mario parte <u>per</u> Roma.</i>	<i>Mario fährt nach Rom.</i>
	Grammatische Präposition	<i>Ti invito <u>alla</u>⁵³ mia festa.</i>	<i>Ich lade dich zu meinem Fest ein.</i>

Tabelle 13: PP-Äquivalente im Englischen und Italienischen: Beispiele.⁵⁴

Tabelle 13 zeigt, dass deutschen Präpositionen und PP im Englischen und im Italienischen formgleiche Äquivalente bzw. ähnliche sprachliche Mittel gegenüberstehen. Eine 1:1-Entsprechung zwischen den Präpositionen bzw. PP im Deutschen und Englischen bzw. Italienischen liegen jedoch nicht immer vor⁵⁵.

1.4 Zusammenfassung

Die folgende Tabelle (Tab. 14) dient als Zusammenfassung der Ausführungen des Kapitels und illustriert die unterschiedlichen Perspektiven auf Präpositionen und PP sowie ausgewählte Charakteristiken adverbialer PP und Objekt-PP anhand von Beispielsätzen. Die Abkürzung *SynF* steht für *Syntaktische Funktion*, *SynV* für *syntaktische Valenz* und *SemV* für *semantische Valenz*.

⁵³ Die Form *alla* ist eine Verschmelzung aus der Präposition *a* und dem definiten Artikel *la*.

⁵⁴ Die Beispiele sind z. T. entnommen aus Huddleston/Pullum (2006: 74, 602) und Schwarze (2011: 304/308).

⁵⁵ Siehe die entsprechenden deutschen Übersetzungen der Beispiele. Im Italienischen kann die Präposition *a* im Deutschen z. B. durch *in* oder *an* wiedergegeben werden.

Beispielsätze		4. Sie arbeitet <u>für</u> Firma X.		
1. Sie öffnete die Flasche <u>im</u> Keller.		5. Sie kommt <u>um</u> Mitternacht an.		
2. Sie sitzt <u>auf</u> der Bank.		6. Sie ist stolz <u>auf</u> ihre Leistungen.		
3. Sie steigt <u>in</u> das Fahrzeug ein.		7. Sie stimmt <u>für</u> Partei X.		
	SynF ⁵⁶	SynV ⁵⁷	SemV ⁵⁸ (sem. Rolle)	Präposition ⁵⁹
1	adverbiale PP	PP-Angabe	PP-Modifikator (Lokalrolle: Position)	semantisiert autonom
2	adverbiale PP	PP-Ergänzung	PP-Argument (Lokalrolle: Position)	semantisiert autonom
3	adverbiale PP	PP-Ergänzung	PP-Argument (Lokalrolle: Direktional)	semantisiert regiert ⁶⁰ (Präpositionsreaktion)
4	adverbiale PP	PP-Angabe	PP-Modifikator (Rezipient/Benefaktiv)	semantisiert autonom
5	adverbiale PP	PP-Ergänzung	PP-Argument (Temporativ)	desemantisiert autonom ⁶¹
6	Objekt-PP	PP-Ergänzung	PP-Argument (Lokalrolle: Ziel)	desemantisiert regiert (Präpositionsreaktion)
7	Objekt-PP	PP-Ergänzung	PP-Argument (Rezipient/Benefaktiv)	semantisiert regiert (Präpositionsreaktion)

Tabelle 14: Präpositionalphrasen: Zusammenfassung der verschiedenen Perspektiven und Analyseebenen.

Ich möchte mit dieser Tabelle folgende Charakteristiken hervorheben, die für die empirischen Untersuchungen in dieser Dissertation von Bedeutung sind:

Präposition (Rektion und semantischer Gehalt)

- Prototypische adverbiale PP enthalten autonome Präpositionen
- Prototypische Objekt-PP enthalten regierte Präpositionen
- Prototypische adverbiale PP enthalten semantisierte Präpositionen

⁵⁶ Für die zentrale Terminologie von SynF s. Kapitel 1.2.2.

⁵⁷ Für die zentrale Terminologie von SynV s. Kapitel 1.2.4.

⁵⁸ Für die zentrale Terminologie von SemV s. Kapitel 1.2.4. Die semantischen Rollen sind hier nach Primus (2012) klassifiziert.

⁵⁹ Für die zentrale Terminologie s. Kapitel 1.1.4.

⁶⁰ Breindl (1989: 32) beschreibt, dass solche adverbiale PP im Hinblick auf die Präpositionsfestigkeit ein Verhalten zeigen, das denen von prototypischen Objekt-PP gleicht.

⁶¹ Die Präposition ist hier autonom, da keine Präpositionsreaktion vorliegt (vgl. *Sie kam am Tag/in der Nacht an*).

- Prototypische Objekt-PP enthalten desemantisierte Präpositionen

Semantische Rollen

- Adverbiale PP und Objekt-PP können ähnliche bzw. gleiche semantische Rollen kodieren.

Form-Funktionsverbindungen

- Adverbiale PP repräsentieren insgesamt zuverlässige Form-Funktionsverbindungen
- Objekt-PP repräsentieren insgesamt weniger zuverlässige Form-Funktionsverbindungen
- Semantisierte Präpositionen (in Objekt-PP) repräsentieren zuverlässige Form-Funktionsverbindungen
- Desemantisierte Präpositionen (in Objekt-PP) repräsentieren weniger zuverlässige Form-Funktionsverbindungen

Weiterhin sind aus kontrastiver Perspektive noch die folgenden zwei zentralen Punkte zu ergänzen, die durch Tabelle 14 nicht abgedeckt werden:

- Einige vom Deutschen typologisch verschiedene Sprachen (TV) verfügen über keine Präpositionen bzw. PP (z. B. Ungarisch und Türkisch). PP-Äquivalente werden prototypisch durch Kasussuffixe und/oder Postpositionen realisiert.
- Einige dem Deutschen typologisch ähnliche Sprachen (TA) verfügen über Präpositionen bzw. PP (z. B. Englisch und Italienisch). Diese ähneln in ihrem Gebrauch deutschen Präpositionen und PP, auch wenn nicht immer 1:1-Entsprechungen vorliegen.

2. Aktuelle theoretische Ansätze zum Zweit- und Fremdspracherwerb

Die Spracherwerbsforschung untersucht „Lern- bzw. Erwerbsprozesse von Erst-, Zweit- und Fremdsprachen und den sich entwickelnden (lerner-)sprachlichen Systemen“ (Hufeisen/Riemer 2010: 738). Der Erwerb einer Zweit- bzw. Fremdsprache ist Gegenstand zahlreicher Theorien. In diesem Kapitel werden zentrale Annahmen aktueller Theorien zum Zweit- und Fremdspracherwerb präsentiert, die für die korpusgestützten Fallstudien in der vorliegenden Dissertation eine zentrale Rolle spielen. Unter aktuellen Annahmen zum Zweit- und Fremdspracherwerb (L2-Erwerb⁶²) zähle ich kognitive Spracherwerbstheorien und beschränke mich dabei auf gebrauchsbasierte Ansätze zum L2-Erwerb. Dieses Kapitel verfolgt nicht den Anspruch, einen umfassenden Überblick zu (aktuellen) Zweit- und Fremdspracherwerbstheorien zu geben bzw. einzelne Modelle im Detail vorzustellen. Für eine umfassende Darstellung diesbezüglich verweise ich z. B. auf Ellis (2008b),

⁶² Ich unterscheide hier nicht weiter zwischen dem Zweit- und Fremdspracherwerb und fasse beide unter dem Begriff L2-Erwerb zusammen.

Lightbrown/Spada (2013), Gass/Selinker (2008), Doughty/Long (2003), Robinson/Ellis (2008), VanPatten/Williams (2015), Helbig et al. (2001) (Teil B.), Krumm et al. (2010) (Teil VIII.).

Kognitive Theorien zum L2-Erwerb nehmen grundlegend an, dass es sich beim L2-Erwerb um einen „Lernprozess [handelt], bei dem die gleichen mentalen Prozesse erforderlich sind wie bei anderen kognitiven Leistungen“ (C. Schmidt 2010: 807). Somit wird davon ausgegangen, dass beim Sprachenlernen allgemeine kognitive Lernmechanismen greifen (vgl. Mitchell/Myles/Marsden 2013: 98, Behrens 2009a/b, Ellis/Wulff 2015). Unter dem Begriff *gebrauchsbasierte Ansätze* werden unterschiedliche Theorien und Ansätze gefasst, sodass man nicht von *der (einen)* gebrauchsbasierten Theorie des Spracherwerbs sprechen kann. Gebrauchsbasierte Ansätze bilden eher eine Art Forschungsrahmen und gehen, im Gegensatz zu generativen/universalgrammatischen Ansätzen davon aus, dass kein sprachspezifisches Modul im Gehirn den Spracherwerb steuert bzw. dass grammatische Strukturen und Regularitäten nicht von Geburt an im Gehirn ‚angelegt‘ sind. Grundlegende Annahmen gebrauchsbasierter Ansätze (vgl. Mitchell/Myles/Marsden 2013: 75) nehmen an, dass

- a) Sprachenlernen auf dem sprachlichen Input, d. h. der Verarbeitung von Sprache im Gebrauch basiert und
- b) Sprachlernende Muster und Regularitäten im Sprachgebrauch aus dem Input ableiten, wobei sie auf allgemeine, kognitive Lernmechanismen zurückgreifen

Im Folgenden gehe ich näher auf die zentralen Konzepte gebrauchsbasierter Ansätze zum L2-Erwerb ein.

2.1 Grundlegende Konzepte gebrauchsbasierter Ansätze zum L2-Erwerb

Zu den zentralen Konzepten gebrauchsbasierter Ansätze zum L2-Erwerb gehören (s. Ellis/Wulff 2015: 75f.):

- Form-Funktionsverbindungen
- Assoziatives Lernen
- Rationale Sprachverarbeitung
- Exemplar-basiertes Lernen
- Emergente Beziehungen und Muster

Ich möchte im Folgenden kurz auf die einzelnen Konzepte eingehen. Im Fokus stehen dabei Form-Funktionsverbindungen und Assoziatives Lernen, da diese Konzepte zentral für meine hypothesenüberprüfenden empirischen Untersuchungen im Lernerkorpus sind. Auf Merkmale des assoziativen Lernens von Form-Funktionsverbindungen im Kontext des L2-Erwerbs gehe ich unter 2.2 näher ein.

2.1.1 Form-Funktionsverbindungen

Form-Funktionsverbindungen (auch *Konstruktionen* genannt⁶³) werden als konventionalisierte Form-Funktionspaare verstanden (Elsen 2014: 226), d. h. als

⁶³ Da meine Dissertation sich nicht an den theoretischen Annahmen der Konstruktionsgrammatik orientiert, möchte ich im Folgenden statt von *Konstruktionen* lediglich von *Form-Funktionsverbindungen* sprechen. Zur Konstruktionsgrammatik s. z. B. den Sammelband von Fischer/Stefanowitsch (2006/2007).

Verbindungen aus Form und Funktion bzw. Bedeutung und gelten als „etabliertes Konzept in der Spracherwerbsforschung“ (Behrens 2009b: 427)⁶⁴. Form-Funktionsverbindungen können unterschiedlich abstrakt sein (vgl. Ellis/Wulff 2015: 76): Unter Form-Funktionsverbindungen fallen einfache Lexeme wie *Hund*, aber auch größere Einheiten wie Phrasen (Elsen 2014: 221) oder abstrakte Repräsentationen von Satzstrukturen wie z. B. Subjekt-Verb-Objekt wie im Satz *Tim küsst Maja* oder idiomatische Ausdrücke wie *ins Gras beißen*. Form-Funktionsverbindungen können jedoch auch unter der Wortebene repräsentiert sein, z. B. im Morphem *-en*, das in der deutschen Sprache den Plural ausdrücken kann (vgl. Ellis/Wulff 2015: 76)⁶⁵. Dabei handelt es sich dann um Verbindungen, die eher eine funktionale Bedeutung ausdrücken (vgl. Ellis/Wulff 2015: 76).

Mit der Annahme von Form-Funktionsverbindungen werden somit Annahmen zu Strukturen einer (zu erlernenden) Sprache gemacht; so wird davon ausgegangen, dass „die Grammatik einer Sprache aus einem strukturierten Netzwerk an Konstruktionen“, also Form-Funktionsverbindungen, besteht (Elsen 2014: 226). Ein Satz wird z. B. verstanden als bestehend aus einer Kombination von unterschiedlich abstrakten Form-Funktionsverbindungen (vgl. Ellis/Wulff 2015: 77). Das linguistische Wissen einer Person besteht demnach aus einem riesigen Depot an Form-Funktionsverbindungen, das durch die Auseinandersetzung mit dem sprachlichen Input auf- und ausgebaut wird. Eine Sprache zu lernen bedeutet dann folglich, Form-Funktionsverbindungen unterschiedlicher Abstraktion zu lernen. Das Lernen von Form-Funktionsverbindungen wird maßgeblich durch das assoziative Lernen gesteuert.

2.1.2 Assoziatives Lernen

In der Neurobiologie wird Lernen definiert als „Aufbau von Neuronenpopulationen im Gehirn“ (Grein 2013: 8). Genauer gesagt ist damit der Aufbau neuronaler Verknüpfungen (=Assoziationen) zwischen Neuronen gemeint (vgl. LdN 2019⁶⁶). Im Kontext gebrauchsbasierter Ansätze zum L2-Erwerb versteht man unter assoziativem Lernen die Verknüpfung von Form und Funktion sprachlicher Strukturen. Die Verknüpfung wird durch bestimmte Charakteristiken ihres Vorkommens im sprachlichen Input gesteuert. Eine wichtige Charakteristik ist die Vorkommensfrequenz, und damit die Erwartbarkeit, mit der eine Form in einer bestimmten Bedeutung bzw. Funktion im sprachlichen Input vorkommt (vgl. Ellis/Wulff 2015: 77). Hört man ein Wort zum ersten Mal, so bildet sich eine Erinnerung an die Merkmale dieses Wortes, also eine Art mentale Repräsentation der Merkmale, heraus. Diese Repräsentation wird zusätzlich begleitet durch eine sogenannte ‚detector unit‘. Man kann sich diese vorstellen als eine Art Kontrollinstanz, die den Input auf die Präsenz der mental repräsentierten (phonologischen oder graphematischen) Merkmale eines Wortes überprüft, und, wenn diese Merkmale vorhanden sind, die Repräsentation des Wortes aktiviert – man

⁶⁴ Behrens (2009b) bezieht sich in ihrem Aufsatz vorrangig auf den Erstspracherwerb. Die Aussage gilt jedoch auch für den L2-Erwerb (vgl. Ellis/Wulff 2015: 75).

⁶⁵ Beispiel: *Mensch+[-en]* bedeutet ‚Mehrzahl‘ oder abstrakt gesehen: Nomen +[Morphem -en] drückt ‚Mehrzahl‘ aus.

⁶⁶ Lexikon der Neurowissenschaft, Artikel *assoziatives Lernen* (<https://www.spektrum.de/lexikon/neurowissenschaft/assoziatives-lernen/964>).

‚erkennt‘ ein Wort, das man bereits im Input angetroffen hat. Dabei spielen nicht nur die Verbindung zwischen Form und Funktion eine Rolle, sondern auch der sprachliche Kontext, in dem die Verbindung auftaucht (Lightbrown/Spada 2013: 111). Je häufiger das Wort nun im Input vorkommt, desto häufiger wird seine mentale Repräsentation aktiviert, d. h. die Assoziation zwischen ‚Merkmal‘ und ‚Wort‘ wird gestärkt. Ähnliches gilt auch für Assoziationen zwischen Form und Funktion dieser Form. Im Erwerb einer Sprache bauen Lernende ein immer stärker werdendes Netzwerk von Assoziationen, d. h. Verbindungen, aus. Jede Assoziation zwischen Form und Funktion wird dabei verbucht – häufige und zuverlässige Assoziationspaare sind demnach leichter und schneller zugänglich bzw. werden leichter und schneller aktiviert. Kempe/MacWhinney (1998: 580) zeigen z. B. für den Kasuserwerb im Deutschen und Russischen, dass dieser stark durch assoziatives Lernen gesteuert wird.

Die Frequenz einer Form-Funktionsverbindung hat demnach einen wichtigen Einfluss auf assoziative Lernprozesse im Kontext des L2-Erwerbs. Es wird davon ausgegangen, dass Form-Funktionsverbindungen bzw. sprachliche Einheiten, die in der Zielsprache frequent sind, früher erworben werden als Einheiten, die eine niedrige Frequenz aufweisen (vgl. auch Ellis 2002, Tschirner 2006). Jedoch ist die Frequenz nicht der allein entscheidende Faktor für den Aufbau von Assoziationsverbindungen. Faktoren wie z. B. Erwerbszeitpunkt, Salienz (d. h. Auffälligkeit der Form), die Relevanz für das Verständnis, sowie bereits bestehende, evtl. konkurrierende Form-Funktionsverbindungen beeinflussen assoziatives Lernen zusätzlich (vgl. Ellis/Wulff 2015: 78). Auf Merkmale des assoziativen Lernens und ihre Bedeutung im L2-Erwerb gehe ich unter 2.2. nochmals näher ein.

2.1.3 Rationale Sprachverarbeitung

Das assoziative Lernen ermöglicht es L2-Lernenden, Sprache rational zu verarbeiten, d. h. ein rationales mentales Modell ihrer Zielsprache aufzubauen, das an ihre sprachlichen Erfahrungen zu einem bestimmten Zeitpunkt angepasst ist bzw. diese widerspiegelt (vgl. Ellis/Wulf 2015: 78f.). Das unbewusst aufgebaute mentale Modell ermöglicht den Lernenden, Form-Funktionsverbindungen im Input zu erkennen und diese gar vorherzusagen. Auf diese Weise trägt das Modell dazu bei, das Verständnis und die Produktion der Zielsprache optimal an den jeweiligen kommunikativen Kontext anzupassen (vgl. Ellis/Wulf 2015: 79).

2.1.4 Exemplar-basiertes Lernen

Ein weiteres Konzept gebrauchsbasierter Ansätze ist das Lernen auf Basis von Mustern bzw. Musterbeispielen (engl. *exemplar*). Konkrete sprachliche Beispiele wie z. B. *Guten Tag* oder *Gute Nacht* werden verarbeitet und mit bereits bekannten Form-Funktionsverbindungen abgeglichen, sodass Muster der Art [*Gut_f* + (Tageszeit)⁶⁷] abgeleitet werden, die wiederum für die Rezeption und Produktion von weiteren Ausdrücken wie *Guten Morgen* genutzt werden können. D. h. dass Sprachlernende von konkreten Beispielen abstrahieren, indem sie strukturell äquivalente bzw. ähnliche Beispiele zusammenfassen und aus diesen abstrakte Muster ableiten (vgl. Ellis/Wulff 2015: 79). Diese Muster sind emergent, das bedeutet, sie

⁶⁷ Das tiefgestellte *f* verweist darauf, dass das Adjektiv entsprechend der folgenden Einheit flektiert wird.

entstehen graduell durch die Verarbeitung des sprachlichen Inputs. Dies geschieht implizit, d. h. es handelt sich um einen unbewussten Prozess im Spracherwerb.

Eine relevante Rolle im L2-Erwerb kommt prototypischen Mustern bzw. Form-Funktionsverbindungen zu, also „the exemplars that are most typical of their categories“ (Ellis/Wulff 2015: 79). Die Prototypikalität ergibt sich aus einer impliziten Analyse der Vorkommnisse von z. B. sprachlichen Formen und ihren Funktionen bzw. Distributionen dieser im sprachlichen Input. Übertragen auf sprachliche Form-Funktionsverbindungen sind diejenigen Form-Funktionsverbindungen prototypisch, die häufig im sprachlichen Input vorkommen und bei denen die Form (zuverlässig) mit einer bestimmten Funktion korreliert. Es kommt also nicht primär darauf an, ob eine Form häufig im Input vertreten ist, sondern dass die Form auch zuverlässig in ihrer Funktion interpretiert werden kann. Das bedeutet im Umkehrschluss, dass im Falle von Prototypikalität eine bestimmte Interpretation bestenfalls durch eine eindeutige Form repräsentiert ist. Es ist jedoch so, dass vor allem die frequenten grammatischen Markierungen in der Regel in beide Richtungen hoch ambig sind (vgl. Ellis/Wulff 2015: 80), so z. B. die Realisierung des deutschen Plurals. Das Konzept Plural kann durch verschiedene Morpheme realisiert werden, z. B. *-en* oder *-e*. Die Morpheme *-en* oder *-e* ihrerseits können jedoch nicht nur den Plural realisieren, sondern auch Flexionssuffixe von Verben (*sie laufen, ich laufe*) darstellen. Das bedeutet, dass die Form *-en* oder *-e* nicht zuverlässig mit der Funktion *Plural* verbunden ist, sowie die Funktion – der Plural – nicht nur durch die Form *-en* oder *-e* realisiert wird (vgl. Ellis/Wulff 2013: 80).

Gebrauchsbasierte Ansätze gehen davon aus, dass zuverlässige Form-Funktionsverbindungen leichter erworben werden als weniger zuverlässige, d. h. je zuverlässiger eine Verbindung, desto prototypischer ist sie und desto leichter ist der Erwerb einer solchen Verbindung (vgl. Ellis/Wulff 2015: 80).

2.1.5 Emergente Relationen und Muster

Im Kontext gebrauchsbasierter Ansätze wird Sprache als komplexes adaptives (d. h. sich ständig optimierendes) System angesehen (vgl. Ellis/Wulff 2015: 81). Sprachliche Muster bzw. Systematiken gelten als emergent, d. h. sie entstehen graduell durch die Verarbeitung sprachlichen Inputs. In diesem Zusammenhang gehen gebrauchsbasierte Ansätze davon aus, dass sprachliche Muster im L2-Erwerb entstehen durch die Interaktion allgemeiner kognitiver Lernmechanismen mit dem sprachlichen Input sowie in der Interaktion mit anderen SprecherInnen in verschiedenen sozialen Kontexten (Ellis/Wulff 2015: 76). Damit wird angenommen, dass sprachlicher Input und sprachliche Interaktionen grundlegend im L2-Erwerb sind und sprachliche Muster durch die Verarbeitung von Sprache im Gebrauch entstehen.

2.1.6 Zwischenfazit I

Gebrauchsbasierte Ansätze zum L2-Erwerb fassen unterschiedliche Spracherwerbtheorien zusammen, teilen jedoch die zentrale Annahme im Zweit- und Fremdspracherwerb (L2-Erwerb), dass kein sprachspezifisches Modul im Gehirn den Spracherwerb steuert, sondern im L2-Erwerb grundlegende Lernmechanismen greifen. In ihren Annahmen zum L2-Erwerb gehen VertreterInnen gebrauchsbasierter Ansätze von fünf zentralen Konzepten aus.

Im Konzept *Form-Funktionsverbindung* (Konzept 1) spiegelt sich die Annahme zum Aufbau einer zu erlernenden Sprache wider: Sprache besteht hiernach aus unterschiedlich abstrakten Form-Funktionsverbindungen. Eine Sprache zu lernen bedeutet demnach, Verbindungen zwischen Formen und deren Funktionen zu erlernen. Dies geschieht durch Mechanismen des assoziativen Lernens, einem weiteren Konzept gebrauchsbasierter Ansätze (Konzept 2). Unter *assoziativem Lernen* versteht man die Verknüpfung sprachlicher Formen mit ihrer Bedeutung bzw. Funktion, also dem Aufbau von Form-Funktionsverbindungen. Lernende bauen im L2-Erwerb ein immer stärker werdendes Netzwerk von Verknüpfungen bzw. Form-Funktionsverbindungen aus. Der Aufbau der Verbindungen wird dabei durch außersprachliche Merkmale wie Frequenz oder den Erwerbszeitpunkt beeinflusst, aber auch bestehende Form-Funktionsverbindungen wirken auf den Erwerb von Form-Funktionsverbindungen ein. Im L2-Erwerb wird weiterhin eine *rationale Sprachverarbeitung* angenommen (Konzept 3). Man geht davon aus, dass Lernende ein rationales mentales Modell der zu erlernenden Sprache aufbauen, das an die sprachlichen Erfahrungen der Lernenden angepasst ist bzw. diese widerspiegelt. Sprachproduktionen von Lernenden bieten somit einen Zugang zum aktuellen mentalen Sprachmodell. Des Weiteren wird angenommen, dass der L2-Erwerb auf Basis von Mustern erfolgt (Konzept 4). Sprachenlernende abstrahieren von konkreten Beispielen im sprachlichen Input, indem sie abstrakte Muster von diesen ableiten. Prototypischen Mustern wird hier ein hoher Stellenwert zugeschrieben. Unter prototypischen Mustern werden Form-Funktionsverbindungen verstanden, die häufig im sprachlichen Input vorkommen und bei denen die Form zuverlässig mit einer bestimmten Funktion korreliert. Hier greifen wieder die Mechanismen des assoziativen Lernens: Für zuverlässige Form-Funktionsverbindungen wird ein schneller bzw. leichter Erwerb angenommen. Im letztgenannten Konzept *Emergente Relationen und Muster* (Konzept 5) spiegeln sich erneut grundlegende Annahmen zur Sprache wider: Sprache gilt als ein komplexes adaptives System, im L2-Erwerb entstehen sprachliche Strukturen aus dem Sprachgebrauch heraus – durch die Interaktion allgemeiner Lernmechanismen mit dem sprachlichen Input sowie durch die sprachliche Interaktion mit anderen SprecherInnen in diversen sozialen Kontexten.

Eine zentrale Annahme gebrauchsbasierter Ansätze bezieht sich auf das Wirken allgemeiner kognitiver Lernmechanismen im L2-Erwerb, wobei das assoziative Lernen eine wichtige Rolle spielt. Im Folgenden werde ich daher näher auf die zentralen Merkmale assoziativen Lernens eingehen und ihre Bedeutung für den L2-Erwerb herausarbeiten.

2.2 Merkmale assoziativen Lernens im Kontext des L2-Erwerbs

Als ein Vertreter gebrauchsbasierter Ansätze zum L2-Erwerb versteht Ellis (2002, 2003, 2006a, 2006b, 2008b) das Erlernen einer Sprache als eine Form des kognitiven und assoziativen Lernens, das durch die Interaktion mit der Umwelt bzw. den sprachlichen Stimuli in dieser Umwelt erfolgt. Als Ergebnis werden sprachliche Muster erlernt und mentale bzw. neuronale Verbindungen zwischen Form und Funktion sprachlicher Einheiten aufgebaut. Welche Verbindungen im Einzelnen zu welcher Zeit im Spracherwerbsprozess aufgebaut werden können, hängt mit den Hauptmerkmalen assoziativen Lernens zusammen, die im Folgenden in Anlehnung an Mitchell/Myles/Marsden (2013: 194ff.) aufgeführt sind:

1. Frequenzbasiertes Lernen

2. Overshadowing und Attention Blocking
3. Statistisches Lernen
4. Lernen von Form-Funktionsverbindungen
5. Implizites Lernen

Im Folgenden möchte ich ausschließlich das Phänomen *Overshadowing und Attention Blocking* in den Vordergrund rücken, da dieses spezifisch im L2-Erwerb greift und u.a. den Einfluss bereits erworbener Sprachen zu erklären versucht. Für die anderen Merkmale assoziativen Lernens siehe beispielsweise Ellis (2002b), Gass/Mackey (2002) (zu 1.), Ellis (2006a), MacWhinney (1987, 2001) (zu 3.), Ellis/Larsen-Freeman (2009) (zu 4.) und DeKeyser (2003) (zu 5.).

2.2.1 Overshadowing und Attention Blocking

Aus der allgemeinen Lernprozessforschung ist bekannt, dass der Aufbau neuen Wissens vom bereits bestehenden Wissen beeinflusst wird und umgekehrt – Erfahrungen, die erlernt werden, interagieren miteinander und beeinflussen sich gegenseitig (Ellis 2006b: 173f.). Aus der Forschung zum assoziativen Lernen gehen zwei allgemeine Mechanismen hervor, die die Aufmerksamkeit auf Sprache steuern: Overshadowing und Attention Blocking (Ellis 2006b: 176). Die Phänomene Overshadowing und Attention Blocking beziehen sich auf die selektive Verarbeitung sprachlicher Stimuli.

Nach Ellis (2006b: 179) sind viele grammatische Form-Funktionsverbindungen nicht nur von geringer Salienz (perzeptive Auffälligkeit), sie können auch für das Verständnis einer Äußerung redundant sein. Wenn ein Merkmal im Input redundant ist und die Bedeutung einer Äußerung ohne dieses Merkmal zufriedenstellend extrahiert werden kann, so kann dieses Merkmal nicht auf die Weise verarbeitet werden, wie es für das Lernen dieses Merkmals notwendig wäre (vgl. Mitchell/Myles/Marsden 2013: 105, s. auch DeKeyser 2005: 18). Zu solch einer nicht zufriedenstellenden Verarbeitung kann es kommen, wenn a) andere Merkmale im Input salienter sind, b) das Weltwissen bereits eine zufriedenstellende Interpretation zulässt (ohne das betroffene Merkmal zu beachten) oder c) der/die Lernende, bedingt durch bereits erworbene Sprache(n), gelernt hat, seine Aufmerksamkeit primär auf andere Merkmale zu lenken, um eine Bedeutung bzw. Funktion einer sprachlichen Form zu rekonstruieren. Verantwortlich für die selektive Verarbeitung sprachlicher Stimuli ist nach Ellis (2006b: 177) ein allgemeiner und sinnvoller Mechanismus des Lernsystems: „It makes good rational sense that evidence that an event is of no predictive value should encourage the learning system to pay less attention to that event in the future“. Das bedeutet, dass das Lernsystem ein für ein Ergebnis (d. h. Interpretation einer sprachlichen Form) nicht relevantes Merkmal nicht erkennt und diesem in Zukunft weniger ‚Aufmerksamkeit‘ bei der sprachlichen Verarbeitung schenkt. Dies ist darauf zurückzuführen, dass das (vermeintlich) redundante Merkmal im sprachlichen Input ‚überschattet‘ (engl. *overshadow*) und folglich nicht als relevant für die Interpretation einer Form gespeichert wird. Dies erschwert jedoch den Erwerb dieses Merkmals und hat Auswirkungen auf die Verarbeitung und Speicherung weiteren sprachlichen Inputs (Ellis 2002: 175, Mitchell/Myles/Marsden 2013: 105). Das Phänomen des ‚Overshadowing‘ kann als Konsequenz zu einer erlernten selektiven Aufmerksamkeit, auch

Attention Blocking (Mitchell/Myles/Marsden 2013: 105) oder *learned attention* (Ellis/Wulff 2015: 82) genannt, führen. Attention Blocking kann die Verarbeitung des zielsprachlichen Inputs und somit den Erwerb bestimmter sprachlicher Merkmale im Input erschweren oder gar verhindern. Attention Blocking tritt z. B. ein, wenn gelernt wurde, dass ein Merkmal X eine bestimmte Funktion Y hat. Für den Lernenden ist es dann schwierig, mit Merkmal X eine neue Funktion Z zu assoziieren. Weiterhin kann es z. B. zu Attention Blocking kommen, wenn eine Funktion zuverlässig mit dem Merkmal X ausgedrückt werden konnte, und im Lernprozess nun ein neues Merkmal Y mit derselben Funktion assoziiert werden muss. Eine bereits ausgeprägte, erlernte Assoziation wird in diesen beiden Fällen den Aufbau einer neuen Assoziation (zunächst) blockieren (vgl. Mitchell/Myles/Marsden 2013: 105).

Nach Ellis (2006b: 179f., s. auch Ellis 2003: 72) haben diese Mechanismen einen besonderen Stellenwert im L2-Erwerb. Im Gegensatz zum L2-Erwerb reagiert das Gehirn beim Erwerb der Erstsprache (L1) aufgrund seiner anfänglichen Plastizität äußerst schnell auf den Input, dem es ausgesetzt wird. In Bezug auf die Sprache wird das Gehirn auf diese Weise für die Repräsentationen der L1 optimiert, die Verarbeitung und Produktion der L1 wird somit automatisiert (Ellis/Wulff 2015: 81f.). Zu Beginn des L2-Erwerbs ist somit i.d.R. bereits ein mentales sprachliches Modell der L1 etabliert und die Plastizität des Gehirns ist nicht mehr in der Form gegeben, wie sie für den L1-Erwerb noch bestand. So wie vergangene Erfahrungen auf neue Erlebnisse wirken und bestehendes Wissen die Aneignung neuen Wissens beeinflusst, so wird auch der L2-Erwerb durch das bereits aufgebaute mentale sprachliche Modell der L1 beeinflusst⁶⁸. Nach Mitchell/Myles/Marsden (2013: 115) gibt es „good evidence for the general notion that what has been learned affects what can be processed“.

Als Beispiel nennen Ellis/Wulff (2015: 81) die Referenz auf Tempus, die sowohl durch Flexion des Verbs erfolgt (*spielte* für die Markierung des Präteritums) als auch durch den Einsatz temporaler Adverbien (z. B. *gestern*). So ist es im Satz *Gestern spielte ich Fußball* überflüssig, zur zeitlichen Verortung des Geschehens auf die Flexion des Verbs zu achten; das Adverb *gestern* ist ein ausreichender Indikator für die zeitliche Referenz; es ist zudem salienter, d. h. perzeptiv auffälliger und zuverlässiger in der Markierung der Zeitreferenz als das Flexionssuffix am Verb (s. auch Ellis 2006b: 179). Dieses Wissen steht Lernenden aus dem L1-Erwerb für den L2-Erwerb bereits zur Verfügung⁶⁹. In der Verarbeitung solcher Äußerungen wie *Gestern spielte ich Fußball* wird der/die Lernende zur Identifizierung der Zeitreferenz dann automatisch auf Adverbien und weniger auf die Verbflexion achten. Hier gilt: „redundant cues are overshadowed because the learners’ L1 experience leads them to look elsewhere for the cues to interpretation“ (Ellis/Wulff 2015: 83). Dies hat jedoch zur Konsequenz, dass der/die Lernende ggf. zielsprachliche Charakteristiken der Verbflexion ‚übersieht‘, was wiederum negative bzw. verzögernde Konsequenzen für den Erwerb der Verbflexion hat⁷⁰. Im Verlauf

⁶⁸ Ellis/Wulff (2015: 89) sprechen diesbezüglich von „L1 attentional biases“.

⁶⁹ An dieser Stelle gehe ich von Lernenden aus, die in ihrer Erstsprache lexikalische Einheiten ähnlich zur deutschen Wortartenkategorie *Adverb* aufweisen, die zur zeitlichen Referenz gebraucht werden können.

⁷⁰ Im L2-Erwerb kann Attention Blocking zur Fossilierung, Verfestigung des Sprachgebrauchs von Lernenden führen, die Entwicklung des mentalen sprachlichen Modells der L2 kann stagnieren.

des Spracherwerbprozesses, wenn vielfältiger zielsprachlicher Input verarbeitet und neue Form-Funktionsverbindungen aufgebaut wurden, können solche kommunikativ redundanten Merkmale ebenfalls kognitiv verarbeitet werden (Mitchell/Myles/Marsden 2013: 115).

Die Phänomene *Overshadowing* und *Attention Blocking* haben somit eine hohe Relevanz im L2-Erwerb, da die L1 bereits ausgebildet ist und Assoziationen zwischen Merkmalen im sprachlichen Input und ihren Funktionen bereits aufgebaut und verfestigt sind. Diese Assoziationen wirken auf unterschiedlichen Abstraktionsebenen und beeinflussen die Verarbeitung und den Erwerb sprachlicher Strukturen der Zielsprache in unterschiedlichen sprachlichen Bereichen (z. B. Grammatik, Lexik, Phonetik) – nach (Ellis 2006b: 175) stellt diese Art von Transfer jedoch ein gängiges Lernphänomen dar und ist damit ein zentrales Charakteristikum des L2-Erwerbs (s. auch Uhlisch 1995: 228). Somit sind auf Basis der Eigenschaften des sprachlichen Inputs in Verbindung mit Prinzipien des assoziativen Lernens Vorhersagen über Schwierigkeitsgrade im Erwerb bestimmter zielsprachlicher Merkmale möglich (Mitchell/Myles/Marsden 2013: 127).

2.2.2 Zwischenfazit II

Ein zentrales Konzept gebrauchsbasierter Ansätze zum L2-Erwerb ist das Wirken allgemeiner kognitiver Lernmechanismen des assoziativen Lernens. Im vorliegenden Kapitel habe ich mich auf zwei zentrale Phänomene des assoziativen Lernens im Kontext des L2-Erwerbs beschränkt: *Overshadowing* und *Attention Blocking*. Beide Phänomene beziehen sich auf die selektive Verarbeitung sprachlicher Stimuli im L2-Erwerb. Der Begriff *Overshadowing* verweist auf das Phänomen, dass für das Verständnis redundante sprachliche Merkmale von salienten (d. h. perzeptiv auffälligen) Merkmalen ‚überschattet‘ werden. Die vermeintlich redundanten Merkmale werden somit nicht ausreichend im mentalen Sprachmodell verarbeitet. In der Folge kommt es zum *Attention Blocking*, also zur selektiven Aufmerksamkeit bei der Verarbeitung sprachlicher Strukturen der L2 (*learned attention*). Die Phänomene *Overshadowing* und *Attention Blocking* haben einen besonderen Stellenwert im L2-Erwerb, da bereits bestehende Form-Funktionsverbindungen aus bereits erworbenen Sprachen (vorrangig der L1) die Verarbeitung sprachlicher Strukturen der L2 beeinflussen. Auf diese Weise ist es möglich, Schwierigkeitsgrade im Erwerb bestimmter zielsprachlicher Strukturen zu prognostizieren.

2.3 Methodologie gebrauchsbasierter Ansätze

Ich möchte an dieser Stelle kurz auf die Methodologie gebrauchsbasierter Ansätze zum L2-Erwerb eingehen und die Verwendung von Produktionsdaten aus einem Lernerkorpus für das Forschungsvorhaben dieser Dissertation motivieren.

Mithilfe expliziten Unterrichts, in dem der Fokus auf sprachliche Formen gelegt wird, Lernende folglich zur expliziten Verarbeitung der Strukturen angeregt werden, kann diesem Umstand jedoch entgegengewirkt werden (Ellis/Wulff 2015: 83), auch wenn angenommen wird, dass Sprachunterricht nicht per se den Spracherwerb steuern kann (Pienemann 1984, 1989). Das Phänomen des *Attention Blocking* scheint demnach entsprechend der Spracherwerbskontexte (natürlich/gesteuert) unterschiedlich ausgeprägt zu sein. Beim natürlichen L2-Erwerb ist davon auszugehen, dass der explizite Fokus auf sprachliche Formen fehlt und somit die sprachliche Entwicklung auf einem niedrigen Stadium stagnieren kann (Ellis/Wulff 2015: 83).

Ein zentrales Konzept gebrauchsbasierter Ansätze zum L2-Erwerb ist die rationale Sprachverarbeitung. Die Annahme zur rationalen Sprachverarbeitung betrifft die Art und Weise, wie Lernende Sprache verarbeiten und inwiefern die Verarbeitung in der sprachlichen Produktion des/der Lernenden widergespiegelt wird. Wenn angenommen wird, dass der sprachliche Input der Lernenden, der erfolgreich verarbeitet wurde, indirekt im mentalen Modell der Lernenden abgebildet ist (Ellis/Wulff 2013: 76) und Lernende das mentale Modell in sprachlichen Produktionen anwenden, so ließe sich z. B. aus der Sprachproduktion von Lernenden auf jenes mentale Modell und damit indirekt auf die für die Lernenden als relevant erwiesenen Charakteristiken des sprachlichen Inputs schließen. Sprachliche Produktionen von Lernenden bieten damit einen indirekten Zugang zu sprachlichen Kompetenzen (Ellis/Barkhuizen 2005: 21). Die Auffassung motiviert somit die Untersuchung von Produktionsdaten von L2-Lernenden für ein besseres Verständnis zu Einflussvariablen im L2-Erwerb.

Des Weiteren sind gebrauchsbasierte Ansätze zum L2-Erwerb kompatibel mit der Methodologie der (Lerner-)Korpuslinguistik (vgl. z. B. Ellis 2003: 95). Wie Ellis/Wulff (2013: 83) diesbezüglich erwähnen, stellt die „[c]orpus-based analysis ... a particularly growing trend across usage-based paradigms“ dar. Nach Ellis/Wulff (2013: 83) eignen sich vor allem große, longitudinale, multimodale, auf unterschiedlichen Ebenen annotierte und frei verfügbare Korpora, die den Sprachgebrauch von Lernenden abbilden, am besten für Untersuchungen im Rahmen gebrauchsbasierter Ansätze – allerdings machen Ellis/Wulff (2013: 83f.) auch implizit deutlich, dass derartige Korpora nur vereinzelt zur Verfügung stehen. Aus den Ausführungen lässt sich jedoch eindeutig ableiten, dass Lernerkorpora, die verschiedene Sprachstufen von Lernenden abbilden, eine geeignete Datengrundlage für Untersuchungen zum L2-Erwerb im Kontext gebrauchsbasierter Ansätze darstellen. Dabei sind qualitative, semi-quantitative und/oder umfassende quantitative Auswertungen von Lernerkorpusdaten nützlich (Ellis/Wulff 2013: 84). Unter semi-quantitativen Auswertungen verstehen Ellis/Wulff (2013: 84) z. B. die Analyse einer repräsentativen Stichprobe an Belegen zum Gebrauch eines bestimmten sprachlichen Phänomens (z. B. Gebrauch spezifischer Morpheme).

Die Datengrundlage der vorliegenden Dissertation – das Lernerkorpus MERLIN, in dem Lernende verschiedener Kompetenzniveaus repräsentiert sind – sowie die semi-quantitativen⁷¹ Auswertungen zum Gebrauch von Präpositionen und PP durch Deutschlernende betten sich folglich gut in die Methodologie gebrauchsbasierter Ansätze zum L2-Erwerb ein.

2.4 Zusammenfassung

Dieses Kapitel präsentierte zentrale Annahmen aktueller gebrauchsbasierter Ansätze zum L2-Erwerb sowie die Methodologie derartiger Ansätze. Gebrauchsbasierte Ansätze zum L2-Erwerb nehmen an, dass Sprachenlernen auf dem sprachlichen Input basiert, d. h. der Verarbeitung von Sprache im Gebrauch. Weiterhin gehen sie davon aus, dass Sprachlernende Muster und Regularitäten im Sprachgebrauch, sogenannte Form-Funktionsverbindungen, aus

⁷¹ In den einzelnen Fallstudien dieser Dissertation spreche ich einfacherweise nur von *quantitativen* Auswertungen.

dem sprachlichen Input extrahieren und abstrakte Muster aus diesen ableiten, wobei sie auf allgemeine, kognitive Lernmechanismen zurückgreifen.

Im Rahmen gebrauchsbasierter Ansätze wird die Zuverlässigkeit von Form-Funktionsverbindungen, und damit entsprechende Merkmale der Zielsprache, für den L2-Erwerb betont. VertreterInnen gebrauchsbasierter Ansätze zum L2-Erwerb gehen davon aus, dass durch die Verarbeitung des sprachlichen Inputs Assoziationen (=Verknüpfungen) zwischen Form und Funktion sprachlicher Einheiten aufgebaut werden. Die Assoziationen werden im Verlauf des Erwerbsprozesses gestärkt, wenn die aufgebauten Form-Funktionsverbindungen wiederkehrend im Input auftauchen. Dabei spielt die Zuverlässigkeit der Form-Funktionsverbindungen eine entscheidende Rolle: Je mehrdeutiger bzw. vielfältiger die Funktion einer sprachlichen Form, desto weniger erfolgreich wird diese erworben; der Erwerb verzögert sich im besten Fall, oder es kommt zur sprachlichen Stagnation. Eindeutige und systematische Form-Funktionsverbindungen begünstigen dahingehend den Erwerb. Insgesamt wird dem sprachlichen Input und den Form-Funktionsverbindungen in diesem Input ein hoher Stellenwert im L2-Erwerb zugeschrieben.

Im Kontext des L2-Erwerbs wirken Mechanismen des assoziativen Lernens: Redundante bzw. wenig saliente (d. h. weniger perzeptiv auffällige) Form-Funktionsverbindungen werden ‚überschattet‘ (*Overshadowing*), die Verarbeitung sprachlicher Strukturen wird durch bereits bestehende Form-Funktionsverbindungen maßgeblich gelenkt (*Attention Blocking*). Zentral wirken dabei bereits ausgebildete Lernprozesse/-mechanismen sowie bestehende Form-Funktionsverbindungen aus der L1. Beide haben einen Einfluss auf die Verarbeitung des zielsprachlichen Inputs und somit auf den Erwerb der L2.

In der Methodologie gebrauchsbasierter Ansätze wird der Auswertung von Produktionsdaten ein hoher Stellenwert zugeschrieben. So fügen sich (Lerner-)Korpora und Analyse- und Auswertungsmethoden aus der Lernerkorpusforschung – wie sie in der vorliegenden Dissertation genutzt werden – gut in den Forschungsrahmen gebrauchsbasierter Ansätze zum L2-Erwerb ein.

3. Lernerkorpusforschung

In diesem Kapitel wird ein Überblick über die Lernerkorpusforschung gegeben. Zum besseren Verständnis werden die Lernerkorpusforschung sowie ihre Schnittstellen zu anderen Forschungsfeldern kurz vorgestellt. Im Anschluss werde ich auf den Begriff *Lernerkorpus* eingehen sowie den Aufbau, die Typologie und die Annotationen in einem Lernerkorpus näher präsentieren und diskutieren. Daraufaufgehend stelle ich die zentralen Analysemethoden der Lernerkorpusforschung vor, die ebenfalls in der vorliegenden Dissertation Anwendung finden. Im letzten Teil des Kapitels werden die einzelnen Schnittstellen der Lernerkorpusforschung ausführlich behandelt, wobei hier ein Schwerpunkt auf der Schnittstelle zur Zweit-/Fremdspracherwerbsforschung liegt.

3.1 Einordnung der Lernerkorpusforschung

Unter dem Begriff *Lernerkorpusforschung* versteht man, vereinfachend gesagt, die Analyse von Daten in einem Lernerkorpus (Granger et al. 2015: 1). Die Lernerkorpusforschung entwickelte sich mit dem Aufkommen erster digitaler Lernerkorpora Ende der 1980er Jahre in Europa als Zweig der Korpuslinguistik⁷² und stellt heute ein eigenständiges Forschungsfeld dar (Granger 2008: 259). Die Korpuslinguistik wird in diversen Einführungen, Aufsätzen und Handbüchern behandelt, auf die im Einzelnen hier nicht eingegangen werden muss (siehe dazu u.a. Lemnitzer/Zinsmeister 2015, Scherer 2014, Keibel et al. 2012, Lüdeling/Kytö 2008). Zentral ist jedoch „die Beschreibung von Äußerungen natürlicher Sprachen, ihrer Elemente und Strukturen, und die darauf aufbauende Theoriebildung auf der Grundlage von Analysen authentischer Texte, die in Korpora zusammengefasst sind“ (Lemnitzer/Zinsmeister 2015: 14). Weiterhin lässt sich festhalten, dass die „[k]orpusbasierte Sprachbeschreibung [...] verschiedenen Zwecken dienen [kann]“ (Lemnitzer/Zinsmeister 2015: 14). Der Grundgedanke, Lerneräußerungen als systematische, digitalisierte und maschinell durchsuchbare Datensammlungen in Form von Korpora zu repräsentieren, bestand darin, den Untersuchungsgegenstand der Korpuslinguistik zu erweitern und damit den Fortschritt in Linguistik und Computertechnologie für die systematische und umfassende Erforschung von Lernaltersprache zu nutzen (Granger 1993: 57). Das Potenzial korpuslinguistischer Forschungen war bereits erkannt worden und so erhoffte man sich ebenfalls Vorteile durch den Aufbau von Lernerkorpora und der Anwendung korpuslinguistischer Werkzeuge und Methoden an der Schnittstelle zur Spracherwerbsforschung (vgl. Granger 2012: 7) – ein Ansatz, den Granger (1994) damals als „revolution in applied linguistics“ bewertete.

Neben der Schnittstelle zur Korpuslinguistik hatte die Lernerkorpusforschung von Beginn an einen Anwendungsbezug im Bereich der Spracherwerbsforschung und des Sprachunterrichts. Auch heute noch liegt die Lernerkorpusforschung an der Schnittstelle von Korpuslinguistik (KL), der Zweit-/Fremdspracherwerbsforschung (Z/FSE) sowie dem Zweit-/Fremdsprachenunterricht (Z/FSU). In jüngster Zeit zählt ebenfalls das Anwendungsgebiet der maschinellen Verarbeitung natürlicher Sprachen (natural language processing, NLP) zu einer weiteren Schnittstelle der Lernerkorpusforschung (Granger et al. 2015: 3). Die Anbindung der Lernerkorpusforschung an diese Schnittstellen wird in Abbildung 5⁷³ veranschaulicht.

⁷² Ich folge der Definition von Lemnitzer/Zinsmeister (2015: 14) und sehe die Korpuslinguistik als eine wissenschaftliche Disziplin an. Scherer (2014: 2) beschreibt die Korpuslinguistik als eine Methode.

⁷³ Die Abbildung illustriert lediglich die Schnittstellen der Lernerkorpusforschung zu anderen Forschungsfeldern. Die Abbildung illustriert nicht, wie die einzelnen hier abgebildeten Schnittstellen jeweils zueinander stehen.



Abbildung 5: Die Lernerkorpusforschung und ihre Schnittstellen: Überblick.

Die interdisziplinäre Ausrichtung der Lernerkorpusforschung (s. Abb. 5) macht ebenfalls deutlich, wie schnell sich dieses Forschungsfeld trotz seines jungen Alters bisher entwickelt hat; Granger et al. (2015: 3) sprechen von der Lernerkorpusforschung als einem „fast-expanding field“. Geprägt wurde die internationale Lernerkorpusforschung maßgeblich von Sylviane Granger und ihrem Team am *Centre for English Corpus Linguistics* an der Universität Louvain, Belgien. Diverse Publikationen und Projekte⁷⁴, ein Handbuch⁷⁵, eine eigens ins Leben gerufene Zeitschrift für Lernerkorpusforschung⁷⁶ sowie die Gründung eines Netzwerks (die Learner Corpus Association, LCA)⁷⁷, die eine regelmäßige Konferenz (die Learner Corpus Research Conference) veranstaltet, wurden von Granger und ihrem Team angestoßen und unterstreichen die rasche Entwicklung und Internationalisierung dieses Forschungsfelds. Trotz dieser Internationalisierung ist die Lernerkorpusforschung jedoch vorwiegend auf die englische Sprache fokussiert. In jüngster Vergangenheit lässt sich jedoch eine Ausweitung auf andere Sprachen, wie z. B. das Deutsche beobachten, was positiv zu bewerten ist (Gilquin 2015: 13). Richtungsweisend für die deutsche Lernerkorpusforschung waren (und sind) hier vor allem die Arbeiten von Anke Lüdeling und ihrem Team von der HU Berlin⁷⁸ sowie von Andrea Abel und ihren MitarbeiterInnen am Institut für Angewandte Sprachforschung von EURAC Research in Bozen, Italien⁷⁹. Das Potenzial der Lernerkorpusforschung für den Bereich

⁷⁴ Das Centre for English Corpus Linguistics an der Universität Louvain bietet eine Übersicht über Publikationen im Bereich der internationalen Lernerkorpusforschung (<https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpus-bibliography.html>) sowie über Lernerkorpus-Projekte (<https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>) (Aug. 2018).

⁷⁵ *The Cambridge Handbook of Learner Corpus Research* ist 2015 bei Cambridge UP erschienen.

⁷⁶ Es handelt sich dabei um das *International Journal of Learner Corpus Research (IJLCR)* (s. auch <https://benjamins.com/catalog/ijlcr>) (Mai 2019).

⁷⁷ Homepage der Learner Corpus Association (LCA): <https://www.learnercorpusassociation.org/> (Mai 2019).

⁷⁸ Siehe <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik> (Mai 2019).

⁷⁹ Siehe <http://www.eurac.edu/de/research/autonomies/commul/Pages/default.aspx> (Mai 2019).

des Deutschen als Zweit- bzw. Fremdsprache (DaZ/DaF) wurde ebenfalls bereits erkannt, der Einsatz von Lernerkorpora in diesem Bereich wird in verschiedenen Publikationen diskutiert (z. B. Fandrych/Tschirner 2007, Walter/Grommes 2008⁸⁰, K. Schmidt 2010, Lüdeling/Walter 2010).

3.2 Das Lernerkorpus: Begriffsbestimmung

Unter einem Korpus verstehe ich nach (Lemnitzer/Zinsmeister 2015: 39)

eine Sammlung schriftlicher oder gesprochener Äußerungen in einer oder mehreren Sprachen. Die Daten des Korpus sind digitalisiert, d. h. auf Rechnern gespeichert und maschinenlesbar. Die Bestandteile des Korpus, die Texte oder Äußerungsfolgen, bestehen aus Daten selbst sowie möglicherweise aus Metadaten, die diese Daten beschreiben, und aus linguistischen Annotationen, die diesen Daten zugeordnet sind.

Ein Lernerkorpus erfüllt alle oben genannten Attribute eines Korpus, mit der Besonderheit, dass die sprachlichen Äußerungen⁸¹ von Lernenden einer Fremd- bzw. Zweitsprache stammen. Folglich verstehe ich in Anlehnung an Granger (2002: 7) unter einem Lernerkorpus eine systematisch zusammengestellte, digital aufbereitete, mit linguistischen Annotationen angereicherte und elektronisch zugängliche Sammlung von authentischen (schriftlichen) Produktionen von L2-Lernenden.

Die Primärdaten in einem Lernerkorpus sind kontinuierliche Textdaten von Sprachenlernenden, die bestenfalls im Rahmen freier Produktionsaufgaben entstanden sind (Granger 2008: 261). Der Anspruch, ein Lernerkorpus enthalte authentische, d. h. natürliche Sprachdaten, kann jedoch nur eingeschränkt erfüllbar sein, da die in einem Lernerkorpus enthaltenen Texte i.d.R. in einer gesteuerten Unterrichtsumgebung entstehen und damit elizitiert sind. Dennoch können freie Textproduktionen in diesem Rahmen als ‚natürlich‘ gelten, da sie die authentische Lernsituation bzw. Unterrichtsaktivitäten widerspiegeln (Granger 2002: 8).

Ein prototypisches Lernerkorpus enthält i.d.R. Texte von L2-Lernenden (Granger 2008: 260), z. B. Personen, die Deutsch als Zweit- oder Fremdsprache lernen. In jüngster Vergangenheit werden jedoch auch Korpora mit Texten von SchülerInnen in ihrer Erstsprache (L1) als Lernerkorpora bezeichnet. In diesem Fall werden SchülerInnen als Lernende verstanden, die ihre muttersprachlichen (Schreib-)Kompetenzen auf- und ausbauen. Ein Beispiel für ein muttersprachliches Lernerkorpus ist das *KoKo*-Korpus (Abel et al. 2014, Abel et al. 2016).

In der Typologie der Sprachkorpora ist ein Lernerkorpus als ein Spezialkorpus⁸² zu klassifizieren, das „nur einen bestimmten Teil einer Sprache, z. B. eine Varietät, abbildet“ (Scherer 2014: 99). Im Falle eines Lernerkorpus ist die abgebildete Varietät die Sprachvarietät

⁸⁰ Im Sammelband von Walter/Grommes (2008) werden zweitspracherwerbtheoretische Zugänge von lernerkorpuslinguistischen Zugängen jedoch (noch) strikt getrennt. Letztere werden von Walter/Grommes (2008: 22) ‚nur‘ der Korpuslinguistik zugeordnet.

⁸¹ Mit *Äußerungen* sind an dieser Stelle sowohl schriftliche als auch mündliche Daten gemeint.

⁸² Auch für Sinclair (1995: 24) ist ein Lernerkorpus ein Spezialkorpus („special corpora“, zitiert in Granger 2002: 18).

von Sprachlernenden, die u.a. durch Systematik und Dynamik gekennzeichnet ist. Diese Sprachvarietät wird auch *Lernersprache* oder *interlanguage* genannt (Ellis/Barkhuizen 2005: 54, Selinker 1972, Näheres zur Lernersprache s. Kap. 3.4.1). In Bezug auf die Repräsentativität von Spezialkorpora merkt Scherer (2014: 28) an, dass diese nicht den Anspruch erheben, repräsentativ für eine Sprache in ihrer Gesamtheit zu sein. Das bedeutet, dass auch ein Lernerkorpus nicht den Anspruch hat, repräsentativ für eine Lernersprache zu sein. Gilquin (2015: 10) dahingegen betont den Anspruch von Lernerkorpora auf Repräsentativität. Nach Gilquin (2015) unterscheidet eben dieses Merkmal ein digitales Lernerkorpus von traditionellen Datensammlungen in der Spracherwerbsforschung (Näheres weiter unten). Der Begriff der *Repräsentativität* ist im Kontext von (Lerner-)Korpora jedoch problematisch, da Repräsentativität die Bekanntheit einer Grundgesamtheit voraussetzt, diese jedoch für Korpusdaten (z. B. sprachlichen Äußerungen von Lernenden) „nicht präzise genug“ bestimmt werden kann (vgl. Lemnitzer/Zinsmeister 2015: 48). Anzunehmen ist, dass Gilquin (2015: 10f.) Repräsentativität in einem anderen, weiteren Sinne versteht: Sie betont die Systematizität der Datensammlung in einem Lernerkorpus, die in der Regel externen Kriterien folgt und somit zur Ausgewogenheit des Korpus beiträgt. Ich folge Gilquins Ansicht, und nehme an, dass ein Lernerkorpus den Anspruch verfolgt, in diesem speziellen Sinne repräsentativ für die Lernersprache zu sein, die es abbildet; gleichzeitig möchte in Bezug auf diesen Repräsentativitäts-Anspruch jedoch eine Einschränkung vornehmen. Wie Granger (2002: 7, s.o.) in ihrer Definition erwähnt, werden die Daten in einem Lernerkorpus „for a particular [...] purpose“, also für einen bestimmten (Forschungs-)Zweck, erhoben und können somit auch nur Repräsentativität in Bezug auf eben diesen Zweck anstreben. Im Falle des Lernerkorpus MERLIN, das als primäre Datengrundlage der vorliegenden Dissertation dient, liegt der Anspruch der Repräsentativität vor allem bei den Kompetenzniveaus, die den Gemeinsamen europäischen Referenzrahmen für Sprachen (GeR) folgen und systematisch für jeden Lernertext im Korpus ermittelt wurden (Abel et al. 2014: 111). Insofern strebt das Lernerkorpus MERLIN in erster Linie an, die einzelnen GeR-Niveaus bestmöglich anhand authentischer Sprachdaten darzustellen.

3.2.1 Korpusaufbau und Metadaten

Wie oben erwähnt, stellen kontinuierliche Textdaten von Sprachenlernenden die Primärdaten in einem Lernerkorpus dar. Um die Systematizität dieser Datensammlung zu gewährleisten, folgt der Aufbau eines Lernerkorpus strikten Designkriterien, d. h. die Datensammlungen werden kontrolliert erhoben und Metadaten systematisch erfasst. Die erfassten Metadaten repräsentieren mögliche Variablen, die für die Ausprägung der Lernersprache von Bedeutung sind. Zur Übersicht sind die Variablen in Abbildung 6 dargestellt.

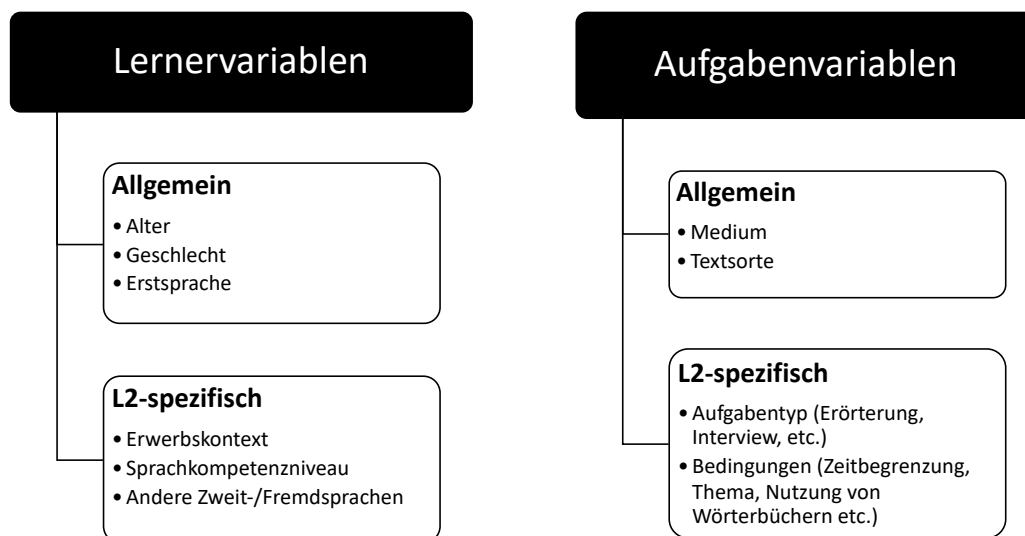


Abbildung 6: Ausgewählte Lerner- und Aufgabenvariablen (Metadaten) zur Dokumentation in einem Lernerkorpus (nach Granger 2008: 264).

Für die Erstellung eines Lernerkorpus ist die Erhebung und Dokumentation diverser Variablen als Metadaten (Abb. 6) zentral. Hierbei ist vor allem die Dokumentation von Lerner- und Aufgabenvariablen

von großer Bedeutung. Die Lernervariablen betreffen die Lernenden (z. B. das Alter, das Geschlecht, die Erstsprache), die Aufgabenvariablen die Merkmale der erhobenen Äußerungen (z. B. Medium oder Textsorte). Speziell in Lernerkorpora findet man L2-spezifische Lernervariablen, z. B. den Erwerbkontext (gesteuert/ungesteuert), das Kompetenzniveau⁸³ sowie L2-spezifische Aufgabenvariablen, z. B. die Bedingungen, unter denen die im Korpus erfassten Daten produziert wurden. Die Variablen werden als Metadaten im Lernerkorpus kodiert. Aufgrund vielfältiger Forschungsinteressen können die kodierten Variablen in ihrer Vielfalt sehr unterschiedlich sein. Die Kontrolle und Dokumentation von Metadaten sind für die spätere Analyse der Lernerdaten und die Interpretation der Ergebnisse essenziell: Werden Lerner- und Aufgabenvariablen dokumentiert, ist es möglich, den Einfluss bestimmter Variablen auf Ausprägungen der Lernaltersprache systematisch zu untersuchen und/oder einzelne Variablen voneinander zu trennen. So kann man in einem Lernerkorpus, in dem das Kompetenzniveau des Lernenden als Variable (Metadatum) dokumentiert ist, diese Variable für die Korpusanalyse berücksichtigen und/oder Teilkorpora mit unterschiedlichen Kompetenzniveaus spezifizieren. Die Kontrolle und Erhebung aller möglichen Einflussvariablen ist jedoch nicht realistisch. Granger (2008: 264) betont, dass eine Darstellung potentieller Variablen, die in einem Lernerkorpus dokumentiert werden können, nie vollständig sein kann, da auch die Variablen, die auf die Lernaltersprache einwirken, sehr vielfältig sind (Granger 2008: 264).

⁸³ Zur Diskussion der Ermittlung von Kompetenzniveaus s. Carlsen (2012) bzw. Kapitel 6 dieser Arbeit.

Die Art der Datenerhebung und -sammlung, und insbesondere die Kontrolle von Variablen in einem Lernerkorpus stellt einen Vorteil dar gegenüber Lernerdaten in der traditionellen Spracherwerbsforschung. Ich gehe weiter unten näher darauf ein.

3.2.2 Typologie

Lernerkorpora können unterschiedlich typologisiert werden. In Bezug auf mögliche Dimensionen orientiere ich mich in diesem Abschnitt vorrangig an Granger (2008) sowie Gilquin (2015)⁸⁴. Die Dimensionen sind in Abbildung 7 zusammengefasst und betreffen die in einem Lernerkorpus enthaltenen Daten, die repräsentierten Lernenden sowie den Kontext, in dem das Lernerkorpus aufgebaut wurde. Die hier vorgestellte Typologie (Abb. 7) erhebt keinen Anspruch auf Vollständigkeit, sondern soll dazu dienen, die zentralen Charakteristika von Lernerkorpora im Allgemeinen zu illustrieren und aktuelle Tendenzen im Aufbau von Lernerkorpora anzuzeigen. Dabei sind in Abbildung 7 die am häufigsten vertretenen Ausprägungen je Dimension fettgedruckt. Im Folgenden gehe ich kurz auf die einzelnen Dimensionen ein.

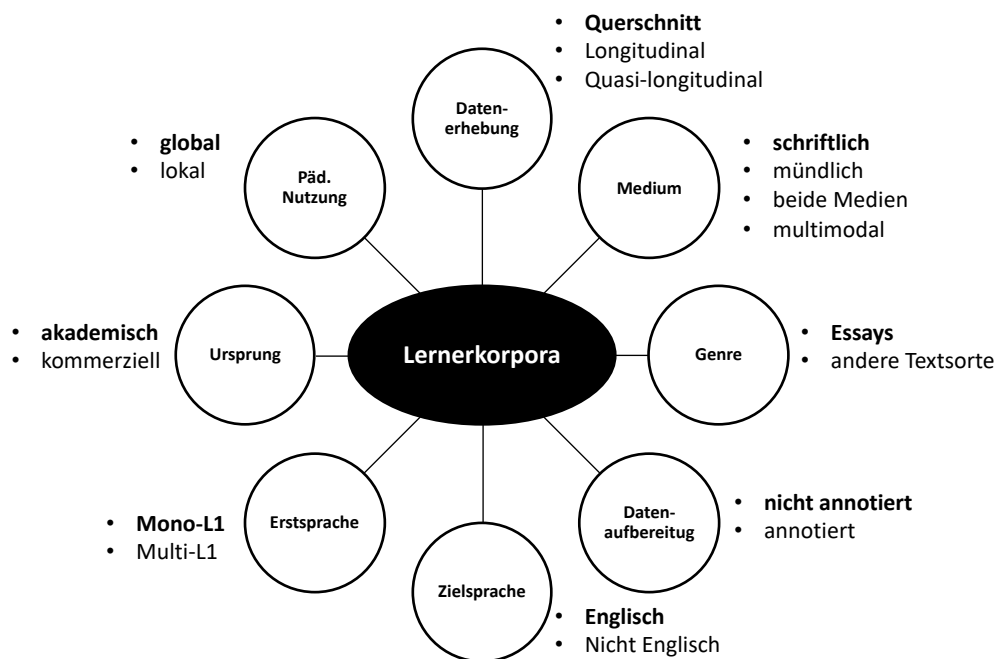


Abbildung 7: Lernerkorpus-Typologie (nach Granger 2008 und Gilquin 2015)

3.2.2.1 Datenerhebung

Die in einem Lernerkorpus enthaltenen Daten unterscheiden sich in Bezug auf die Art der Datenerhebung. Querschnitts-Lernerkorpora enthalten Daten von einer Gruppe von Lernenden, die zu einem einzigen Zeitpunkt erhoben werden. Longitudinale Lernerkorpora beinhalten Daten von ein- und denselben Lernenden, die über einen bestimmten Zeitraum zu mehreren Zeitpunkten erhoben werden und sich daher für Untersuchungen zum Sprachentwicklungsverlauf eignen. Aufgrund des zeitlichen Aufwands einer longitudinalen Datenerhebung sind entsprechende Lernerkorpora jedoch sehr selten. Abhilfe bieten jedoch

⁸⁴ Ähnliche Klassifizierungen finden sich ebenfalls in Nesselhauf (2004) sowie Granger (2002, 2012).

quasi-longitudinale Lernerkorpora, auch pseudo-longitudinale Lernerkorpora genannt (Gilquin 2015: 14). Quasi-longitudinale Lernerkorpora enthalten Daten, die zu einem einzigen Zeitpunkt erhoben werden, jedoch von mehreren Gruppen von Lernenden mit jeweils unterschiedlichen Kompetenzniveaus. Die Kompetenzniveaus gilt es dann jedoch möglichst systematisch zu bestimmen (Gilquin 2015: 14).

3.2.2.2 Medium

Die meisten Lernerkorpora enthalten schriftliche Texte von Sprachlernenden. Lernerkorpora, die ausschließlich oder zusätzlich die gesprochene Sprache (Tonaufnahmen und/oder Transkriptionen) abbilden, sind vergleichsweise selten. In jüngster Vergangenheit wurden ebenfalls multimodale Lernerkorpora aufgebaut, die z. B. Ton- und Videoaufnahmen von Sprachlernenden kombinieren (Gilquin 2015: 12).

3.2.2.3 Genre

Bei den in einem (schriftlichen) Lernerkorpus vertretenen Textsorten dominieren Essays, d. h. argumentative Texte, die zumeist in einem universitären, akademischen Kontext verfasst sind (Granger 2008: 262). Sehr interessant für die Forschung, jedoch auch besonders selten, sind diejenigen (wenigen) Lernerkorpora, die Texte unterschiedlicher Textformen von ein und denselben Lernenden enthalten. Derartige Lernerkorpora ermöglichen den Vergleich von Lerner Sprache über verschiedene Textgenres hinweg (Gilquin 2015: 13).

3.2.2.4 Datenaufbereitung

Lernerkorpora unterscheiden sich ebenfalls in der Aufbereitung der Primärdaten. Die meisten Lernerkorpora enthalten neben den Primärdaten, den Lernertexten, und den Metadaten keine weiteren Informationen in Form linguistischer Annotationen (Granger 2008: 265)⁸⁵. Dies hängt u.a. damit zusammen, dass standardisierte Annotationswerkzeuge auf Basis von redigierten, sprachlich konformen Texten trainiert und entwickelt werden, die aber nur bedingt für die automatische Annotation von Lernertexten geeignet sind (vgl. z. B. Meurers 2015: 538) bzw. wenn eingesetzt, einen hohen manuellen Korrekturaufwand erfordern. An der Schnittstelle zur automatischen Sprachverarbeitung (NLP) werden jedoch bereits Lösungsansätze hierzu erforscht, um die automatische Annotation von Lernerkorpora in verschiedener Hinsicht zu optimieren (van Rooy 2015: 79). Diejenigen Lernerkorpora, die über linguistische Annotationen verfügen, enthalten in der Regel Lemmatisierungen (Annotationen der Grundform einzelner Wortformen) sowie Wortartenannotationen.

Einige wenige Lernerkorpora verfügen zusätzlich über manuelle Annotationen wie z. B. Zielhypothesen (Lüdeling 2008) oder Fehlerannotationen (Granger 2008: 266, siehe auch Lüdeling/Hirschmann 2015). Zielhypothesen sind Rekonstruktionen (d. h. Korrekturen) von Lerneräußerungen in einem Korpus, und dienen als Basis für weitere Annotationsebenen. Fehlerannotationen sind, vereinfacht dargestellt, Auszeichnungen von Fehlern mit vorab festgelegten und klar definierten Fehlerkategorien. Fehlerannotationen zählen zu den wichtigsten Annotationen in Lernerkorpora (Díaz-Negrillo/Thompson 2013: 13). Im Gegensatz zu automatischen Annotationen, bei denen standardisierte Werkzeuge und standardisierte

⁸⁵ Bei Gilquin (2015) findet man zu Annotationen in Lernerkorpora keine spezifischen Angaben.

Kategorien bereits vorliegen (und ggf. adaptiert werden können), fehlt bisher ein entsprechender Standard im Bereich der Fehlerannotationen. Automatische sowie manuelle Annotationen sind für die Analyse von Lernerkorpora zentral, sie können jedoch komplex sein und sind mit unterschiedlichen Herausforderungen verbunden. Ich gehe unter Kapitel 3.3. ausführlich auf Annotationen in Lernerkorpora ein.

3.2.2.5 Zielsprache

Lernerkorpora mit Äußerungen von Lernenden des Englischen als Zielsprache dominieren die Lernerkorpus-Landschaft (Granger 2008: 262), die Zahl an Lernerkorpora anderer Zielsprachen (z. B. Deutsch) wächst jedoch stetig an. Während Anfang der 2000er Jahre Nesselhauf (2004) noch verkündet, dass nahezu alle existierenden Lernerkorpora solche der Zielsprache Englisch sind, erwähnen Granger (2008: 262) und Gilquin (2015: 13) nur einige Jahre später bereits Lernerkorpora der Zielsprachen Französisch, Schwedisch, Norwegisch, Niederländisch, Spanisch, und auch Deutsch. Eine Übersicht zu aktuell bekannten Lernerkorpora unterschiedlicher Zielsprachen bietet die Website *Learner corpora around the world (LCW)* des Centre for English Corpus Linguistics an der Universität Louvain⁸⁶.

3.2.2.6 Erstsprache

Unter den bisher bekannten Lernerkorpora dominieren solche, die Daten von Lernenden mit einer einzigen Erstsprache (L1) enthalten. Solche Lernerkorpora werden auch als mono-L1 klassifiziert. Nach Gilquin (2015: 13) dominieren hier Lernende mit asiatischen Erstsprachen. Multi-L1 Lernerkorpora verfügen dahingegen über Daten von Lernenden unterschiedlicher Erstsprachen. Multi-L1-Lernerkorpora eignen sich vor allem für die Erforschung des Einflusses einer Erstsprache auf den Gebrauch der Zielsprache, vorausgesetzt die produzierten Daten sind gut miteinander vergleichbar (Gilquin 2015: 13f.). Zu beachten ist auch, dass nicht in allen Lernerkorpora die L1 der Lernenden dokumentiert ist (Gilquin 2015: 14).

3.2.2.7 Ursprung

Lernerkorpora haben ihren Ursprung hauptsächlich in einem akademischen Kontext (z. B. im Rahmen eines Forschungsprojekts), werden i.d.R. zu wissenschaftlichen Forschungszwecken aufgebaut und sind für die wissenschaftliche Gemeinschaft zugänglich. Diese so genannten akademischen Lernerkorpora unterscheiden sich von kommerziellen Lernerkorpora, die z. B. von Verlagsunternehmen aufgebaut werden und für die wissenschaftliche Gemeinschaft nicht zur Verfügung stehen (Granger 2008: 261). Kommerzielle Lernerkorpora sind zudem tendenziell multi-L1-Lernerkorpora während akademische Lernerkorpora eher mono-L1-Lernerkorpora umfassen (Granger 2012: 12).

Neben der Unterscheidung zwischen akademischen und kommerziellen Lernerkorpora möchte ich zusätzlich unterscheiden zwischen frei verfügbaren und nicht frei verfügbaren Lernerkorpora. Die Dimension der freien Verfügbarkeit wird weder bei Granger (2008) noch bei Gilquin (2015) erwähnt. Sind kommerzielle Lernerkorpora generell nicht frei verfügbar, unterscheiden sich akademische Lernerkorpora hinsichtlich ihrer freien (d. h. kostenfreien)

⁸⁶<https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html> (Mai 2019).

Verfügbarkeit: Eines der größten und bekanntesten Lernerkorpora mit Englisch als Zielsprache, das International Corpus of Learner English (ICLE), ist ausschließlich gegen Entgelt⁸⁷ zugänglich, während andere Lernerkorpora wie z. B. Falko⁸⁸ (Lüdeling et al. 2008) oder MERLIN (Abel et al. 2014) frei und unentgeltlich für die Öffentlichkeit zur Verfügung stehen.

3.2.2.8 Pädagogische Nutzung

Im Hinblick auf den Einsatz von Lernerkorpora kann man nach Granger (2008: 263) zunächst zwischen einer direkten und indirekten pädagogischen Nutzung von Lernerkorpora bzw. zwischen lokalen und globalen Lernerkorpora (Gilquin 2015: 15) unterscheiden. Ein lokales Lernerkorpus wird im Rahmen regulärer Unterrichtsaktivitäten (durch die Lehrperson) erhoben, um im Anschluss wieder (nur) für die Unterrichtspraxis genutzt zu werden. Bei lokalen Lernerkorpora sind die Lernenden somit gleichzeitig ProduzentInnen und NutzerInnen der Korpusdaten (daher: direkte pädagogische Nutzung). Globale Lernerkorpora dahingegen werden im Rahmen von Forschungsprojekten aufgebaut und umfassen dementsprechend Daten einer größeren Gruppe von Lernenden. Anders als bei lokalen Lernerkorpora ist bei globalen Lernerkorpora die pädagogische Nutzung nicht vorab festgelegt, jedoch können die Lernerdaten bzw. deren Analyse auch für die pädagogische Nutzung eingesetzt werden (daher: indirekte pädagogische Nutzung). Die meisten Lernerkorpora zählen zu den globalen Lernerkorpora (Gilquin 2015: 15).

Die obigen Dimensionen und ihre Ausprägungen spiegeln die zentralen Charakteristika von Lernerkorpora im Allgemeinen wider. In der vorliegenden Dissertation bildet das Lernerkorpus MERLIN die Datengrundlage. Das Lernerkorpus wird in Kapitel 6 vorgestellt und entsprechend der hier vorgestellten Typologie eingeordnet.

3.3 Annotationen in einem Lernerkorpus

Nach Granger (2008: 265) verfügen die meisten Lernerkorpora über keine Annotationen. Ich gehe an dieser Stelle jedoch auf Annotationen in Lernerkorpora ein, da das Lernerkorpus MERLIN, das als Datengrundlage dieser Dissertation dient, über verschiedene Annotationen verfügt. Bei den Ausführungen in diesem Kapitel gehe ich jedoch nicht spezifisch auf MERLIN ein (vgl. dazu Kap. 6), sondern nehme an dieser Stelle die zentralen Annotationen in Lernerkorpora der Zielsprache Deutsch in den Fokus.

Annotationen sind „kontextualisierte Analysen“ von Daten (Lemnitzer/Zinsmeister 2015: 57). In einem Korpus werden die Daten durch Annotationen mit weiteren linguistischen Informationen (z. B. zu Wortarten) angereichert. Da es sich bei Annotationen um zusätzliche Informationen bzw. Analysen handelt, sind diese auch immer interpretativ (van Rooy 2015:

⁸⁷ Die aktuelle Version des Lernerkorpus ICLE (Granger et al. 2009) kann aktuell nur als CD-ROM erworben werden. Der Preis beträgt 272,25€. (siehe I6doc 2019).

⁸⁸ Falko steht für *fehlerannotiertes Lernerkorpus*. Das Lernerkorpus wurde an der Humboldt Universität zu Berlin aufgebaut. Näheres zum Aufbau des Lernerkorpus s. Reznicek et al. (2012) und <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/standardseite> (Mai 2019).

79, Lüdeling/Walter 2009: 23). Annotationen sind damit auch immer interpretative Informationen, die in einem Korpus hinzugefügt werden. Der Zweck von Annotationen liegt darin, eben diese linguistischen Informationen gezielt und effizient extrahieren zu können, um sie z. B. weiteren Analysen zugänglich zu machen. Somit dienen „Annotationen als eine Art Anker, und [können] dem Nutzer ermöglichen, auf effiziente Weise relevante Beispiele in einem Korpus zu finden“ (Lemnitzer/Zinsmeister 2015: 57). Lüdeling/Hirschmann/Shadrova (2017: 13) sehen in Annotationen „one of the biggest breakthroughs in empirical work“, da auf diese Weise die Kategorisierungen, die linguistischen Analysen zugrunde liegen, explizit und nachvollziehbar gemacht werden können.

Im vorliegenden Kapitel unterscheidet ich zwischen automatischen und manuellen Annotationen. Zu den automatischen Annotationen zähle ich solche Annotationen, die mithilfe standardisierter computerlinguistischer Werkzeuge (Tools) auf Basis von „Regeln oder aus Korpora abgeleiteten Wahrscheinlichkeiten“ (Lemnitzer/Zinsmeister 2015: 60) automatisch erzeugt werden. Hierzu zählen z. B. die Wortartenannotationen (auch Part-of-Speech-Tagging, kurz POS-Tagging). Unter manuellen Annotationen verstehe ich dahingehend ‚in Handarbeit‘ durchgeführte Annotationen, d. h. die Anreicherung der Daten mit den entsprechenden Informationen erfolgt durch eine Person. Hierbei kann auf Editierungsprogramme zurückgegriffen werden, die als Unterstützung bei der manuellen Annotation dienen und diese somit beschleunigen können. Zu den manuellen Annotationen zählen z. B. die Auszeichnung von Abweichungen in einem Lernerkorpus (Fehlerannotationen)⁸⁹. Im Folgenden möchte ich mich auf Annotationen in Lernerkorpora beschränken und dabei sowohl auf automatische als auch auf manuelle Annotationen eingehen. Bei den automatischen Annotationen liegt der Schwerpunkt auf Wortartenannotationen, bei den manuellen Annotationen werden Zielhypothesen und Fehlerannotationen zentral behandelt. Abschließend gehe ich auf die Repräsentation von Annotationen in der Lernerkorpusarchitektur ein.

3.3.1 Automatische Annotationen: Wortarten

Die Anwendung computerlinguistischer Werkzeuge zur automatischen Annotation lernersprachlicher Daten in einem Lernerkorpus ist mit vielen Herausforderungen verbunden. Die verfügbaren computerlinguistischen Werkzeuge werden auf Basis von Textkorpora, in der Regel Zeitungstexten, entwickelt und auf diesen Daten trainiert⁹⁰. Die Qualität der

⁸⁹ Bei anderen Autoren wird zwischen automatischen, semi-automatischen und manuellen Annotationen unterschieden. Van Rooy (2015: 87) nennt semi-automatische Annotation auch *interaktive Annotationen* und versteht darunter eine Kombination aus automatischer Annotation mit anschließender manueller Korrektur; Lemnitzer/Zinsmeister (2015: 60) sprechen in diesem Zusammenhang lediglich von interaktiven Annotationsprogrammen, mit deren Hilfe automatische Annotationen manuell angepasst werden können. Da der Annotationsvorgang im Kern automatisch bleibt, zähle ich eine Kombination aus automatischer Annotation mit anschließender manueller Korrektur zu den automatischen Annotationen. Für Granger (2002: 16) sind semi-automatische Annotationen dahingegen manuelle Annotationen, die mithilfe eines Editierungsprogramms (z. B. dem Programm MMAX2, Müller/Strube 2006) in einen Text eingefügt werden.

⁹⁰ Ich möchte an dieser Stelle nicht näher auf die Verfahrensweisen automatischer Annotationswerkzeuge eingehen. Für einen Überblick siehe z. B. Lemnitzer/Zinsmeister (2015: 68ff.),

Annotationen variiert dementsprechend u.a. in Abhängigkeit davon, wie stark sich die Trainingsdaten von den zu annotierenden Daten unterscheiden (Lüdeling/Walter 2009: 22). Aus diesem Grund arbeiten „[a]utomatische Tools [...] am besten auf dem Datentyp, mit dem sie entwickelt wurden“ (Zinsmeister 2015: 102). Insgesamt erreichen automatische Annotationsmethoden eine sehr hohe Genauigkeit (mehr als 95%⁹¹), sodass manuelle Korrekturen nur in wenigen Fällen notwendig sind (Schmid 1994, 2008).

Lernerkorpora enthalten Daten, die in vielen Fällen von der zielsprachlichen Norm abweichen, so z. B. durch non-konforme Schreibweisen und andere Abweichungen, die eine automatische Analyse bzw. Annotation erschweren (van Rooy 2015: 87, Granger 2008: 265, Zinsmeister 2015: 107). Diese Abweichungen sind für eine Beschreibung von Lernaltersprache von hohem Interesse (vgl. z. B. van Rooy 2015: 80), können jedoch eine Herausforderung für automatische Annotationswerkzeuge darstellen. Am Beispiel der automatischen Wortartenannotation soll diese Herausforderungen im Folgenden exemplarisch illustriert werden.

Für deutsche Korpora hat sich das Stuttgart-Tübingen-Tagset (STTS)⁹² (Schiller et al. 1999) als Standard für die Wortartenannotation etabliert. Die Wortartenklassifikation des Tagsets basiert auf lexikalischen, morphologischen, semantischen und syntaktischen Kriterien (Zinsmeister 2015: 107, Lemnitzer/Zinsmeister 2015: 63f.). Die lexikalischen und syntaktischen Kriterien für die Zuweisung eines Wortartentags nach STTS werden statistisch auf Basis der Trainingsdaten ermittelt. Wie oben bereits erwähnt, handelt es bei den Trainingsdaten in der Regel um Zeitungstexte. In prototypischen Fällen sind lexikalische und syntaktische Kriterien kompatibel, das entsprechende Wortartentag kann eindeutig ermittelt werden. In Einzelfällen können sich die Kriterien (auch in zielsprachlich normkonformen Texten) jedoch widersprechen z. B. *ohne Wenn und Aber*, in denen die Wortformen *Wenn* und *Aber* lexikalisch als Konjunktionen klassifiziert werden, syntaktisch jedoch als Nomen zu bestimmen sind (Reznicek/Zinsmeister 2013: 67). Solche Einzelfälle können dann lediglich in einem zweiten automatischen Analyseschritt disambiguiert werden (Reznicek/Zinsmeister 2013: 67).

Abweichungen in Lernertexten können ebenfalls Wortformen enthalten, die hinsichtlich der lexikalischen, morphologischen oder syntaktischen Merkmale eine eindeutige Zuweisung von Wortarten-Tags erschweren, wie folgendes Beispiel aus Reznicek/Zinsmeister (2013: 64) illustriert:

[...] *in der Heutzutager Gesellschaft*. (FalkoEssayL2v2.4).

speziell für POS-Annotationen s. Schmid (1994, 2008), für morphologische Annotationen Schmid/Fitschen/Heid (2004), Fitschen/Gupta (2008).

⁹¹ Schmid (1994) berichtet von 95%-96% Genauigkeit bei der Zuweisung von Wortarten auf Basis eines stochastischen Taggers (TreeTagger).

⁹² Einen Überblick über das Tagset erhält man unter: <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-table.html> (Mai 2019).

Der Stamm der Wortform *heutzutage* legt eine Klassifikation als Adverb nahe, dahingegen sprechen die Großschreibung sowie das Derivationsuffix *-er* für eine Klassifikation als Nomen. Die syntaktische Information zur Position zwischen Artikel (*der*) und Nomen (*Gesellschaft*) deutet jedoch auf ein attributives Adjektiv hin. Die automatische Wortartenzuweisung, die primär auf lexikalische und morphologische Merkmale zurückgreift, ordnet für das Wort *Heutzutage* das Wortarten-Tag *NN* zu und klassifiziert diese Wortform somit als Nomen (Reznicek/Zinsmeister 2013: 64).

Verschiedene Ansätze wurden entwickelt, um mit der Problematik der automatischen Annotation von Lernerkorpusdaten umzugehen. Eine mögliche Lösung wäre ganz auf automatische Annotationen in Lernerkorpora zu verzichten. Diese Lösung ist zwar naheliegend, jedoch verzichtet man somit ebenfalls auf wertvolle, zusätzliche Informationen, die bei der Korpusrecherche und -analyse hilfreich sein können (van Rooy 2015: 83). Neue Ansätze zur automatischen Annotation von Lernerkorpusdaten versuchen daher, die automatischen Annotationswerkzeuge für Lernerdaten zu optimieren und z. B. die für einen automatischen Tagger problematischen Einheiten wie die Wortform *Heutzutage* im obigen Beispiel zur Erforschung von Lernerdsprache zu nutzen. Derartige Ansätze präsentieren Reznicek/Zinsmeister (2013), Díaz-Negrillo et al. (2010) sowie Rastelli (2009) für die automatische Wortartenannotationen in Lernerkorpora. Rehbein et al. (2012), Dickinson/Ragheb (2009) und Ragheb/Dickinson (2013) zeigen einen möglichen Zugang zur automatischen Annotation syntaktischer Strukturen (auch *Parsing* genannt) von Lernerkorpusdaten.

Ansätze, die sich mit dem Nutzen automatischer Annotationswerkzeuge für Lernerkorpora beschäftigen sind relativ neu (Meurers 2015: 545) und fokussieren hauptsächlich Lernerkorpora mit Englisch als Zielsprache. Rehbein et al. (2012) sowie Reznicek/Zinsmeister (2013) zeigen Lösungsansätze für die automatische Annotation deutscher Lernerkorpora auf. An dieser Stelle gehe ich nicht näher auf diese Ansätze ein, da sie für das Forschungsvorhaben der vorliegenden Dissertation nicht zentral sind. Das Lernerkorpus MERLIN, das als Datengrundlage dieser Dissertation dient, enthält Wortartenannotationen, die auf Basis traditioneller Annotationswerkzeuge für die deutsche Sprache erzeugt wurden.

3.3.2 Manuelle Annotationen: Zielhypothese und Fehlerannotation

Manuelle Annotationen spielen in der Lernerkorpusforschung eine wichtige Rolle. In diesem Abschnitt möchte ich auf zwei zentrale Ebenen der manuellen Annotationen näher eingehen und ihre Relevanz in der Lernerkorpusforschung verdeutlichen: 1) Die Annotationen einer Zielhypothese (ZH), auch Target Hypothesis (TH) genannt, und 2) die Fehlerannotation, also die Annotationen nicht-zielsprachlicher Strukturen. Zu weiteren manuellen Annotationen in deutschen Lernerkorpora (z. B. topologische Felder, Satzspannen oder Aspekte der Höflichkeit) sei auf Lüdeling et al. (2005), Reznicek et al. (2012) und Wisniewski et al. (2014) verwiesen.

3.3.2.1 Zielhypothesen

Die Zielhypothese (ZH) stellt eine Art Normalisierungsebene im Lernerkorpus dar (Lemnitzer/Zinsmeister 2015: 87). Die ZH bildet eine zielsprachliche Struktur ab und ermöglicht auf diese Weise z. B. eine weitere automatische Verarbeitung der im Korpus enthaltenen Daten. Die ZH ist somit zu verstehen als die zielsprachliche Rekonstruktion einer (fehlerhaften) Lerneräußerung⁹³ (Lüdeling 2008: 126). Die Zielhypothese stellt eine Annotationsebene und damit eine Interpretation der Original-Lerneräußerung dar. Wie die Umschreibung ‚zielsprachliche Rekonstruktion‘ vermuten lässt, ist die ZH vor allem für die Fehlerannotation in Lernerkorpora zentral, da sie die Auszeichnung von Fehlern erst nachvollziehbar und transparent macht. Ich gehe weiter unten näher auf Fehlerannotationen ein. Abbildung 8 gibt ein Beispiel für eine ZH, Abweichungen zur Original-Lerneräußerung sind fettgedruckt. Im Beispiel wurde die Flexion des Adjektivs *neu* verändert sowie die Schreibweise des Adjektivs *offen* korrigiert.

Original-Lerneräußerung (O)	Ich bin flexibel und für neuen Aufgaben offen.
Zielhypothese (ZH)	Ich bin flexibel und für neue Aufgaben offen .

Abbildung 8: Beispiel für eine Zielhypothese (ZH). Entnommen aus dem Lernerkorpus MERLIN (ID: 1023_0001416).

Bei der Repräsentation der ZH als Annotationsebene in einem Lernerkorpus kann man zwischen verschiedenen Varianten bzw. Formen der ZH unterscheiden: Im Lernerkorpus Falko sind das beispielsweise die ZH0, die ZH1 und die ZH2 (s. Reznicek et al. 2012: 38f., Reznicek/Zinsmeister 2013: 69). Die ZH0 gilt als „vereinfachte Variante der Zielhypothese“ (Reznicek/Zinsmeister 2013: 69) und enthält keine Bewegungen von Elementen der Original-Lerneräußerung. Die ZH1 wird auch minimale Zielhypothese genannt und enthält lediglich Rekonstruktion auf Ebene der Orthographie, Morphologie und Syntax (Grammatik). Liegen keine Bewegungen von Elementen der Original-Lerneräußerungen auf Ebene der ZH1 vor, sind ZH0 und ZH1 identisch. Die ZH2 enthält dahingegen u.a. Rekonstruktionen lexikalischer und pragmatischer Natur und wird daher auch erweiterte Zielhypothese genannt (vgl. Reznicek et al. 2012: 38, MERLIN project 2014: 15). Die folgende Abbildung (Abb. 9) illustriert die drei ZH-Varianten ZH0, ZH1 und ZH2 anhand eines Beispiels.

O	Diese	Erfahrung	wurde	mich	späte	gut	helfen	,	bei	[...]	[...]
ZH0	Diese	Erfahrung	würde	mir	später	gut	helfen	,	bei	[...]	[...]
ZH1	Diese	Erfahrung	würde	mir	später	gut	helfen	,	bei	[...]	[...]
ZH2	Diese	Erfahrung	würde	mir	später	sehr			bei	[...]	helfen

Abbildung 9: Beispiel für die ZH-Varianten ZH0, ZH1 und ZH2 am Beispiel einer Lerneräußerung aus MERLIN (ID: 1023_0001418)⁹⁴.

⁹³ Es sei an dieser Stelle betont, dass es sich bei der ZH um keine Rekonstruktion der Lernerintention handelt, da diese anhand des schriftlichen Textes nicht rekonstruierbar ist (Lüdeling/Hirschmann 2015: 141).

⁹⁴ ZH0 und ZH1 sind identisch, da auf ZH1 keine Bewegung von Elementen aus O (=Original-Lerneräußerung) vorgenommen wurde.

Im Lernerkorpus MERLIN, das als primäre Datengrundlage dieser Dissertation dient, finden sich minimale (ZH1) sowie erweiterte Zielhypothesen (ZH2). Im Folgenden werde ich die Abkürzung ZH verwenden, um allgemein auf die Annotationsebene *Zielhypothese* zu verweisen. Die Abkürzungen ZH0, ZH1 und ZH2 verwende ich, wenn ich explizit auf die jeweilige Annotationsebene referieren möchte.

Die Trennung einzelner Varianten der ZH (vgl. Abb. 9) dient u.a. dazu, die Abweichungen zwischen Original-Lerneräußerung und der zielsprachlichen Struktur auf den entsprechenden Ebenen (z. B. Syntax und Pragmatik) möglichst präzise abzubilden. Dies ist nicht nur für die darauf aufbauende Fehlerannotation hilfreich (Lüdeling/Hirschmann 2015: 145), sondern gewährleistet auch eine bessere Evaluation der erstellten ZH-Annotationen: So ergeben sich bei Abweichungen auf orthographischer oder syntaktischer Ebene i.d.R. weniger Interpretationsspielräume und damit eine höhere Konsistenz zwischen unterschiedlichen AnnotatorInnen als bei Erstellung von ZH auf pragmatischer oder lexikalischer Ebene (vgl. Reznicek et al. 2012: 38).

Die ZH bildet nicht nur die Basis für Fehlerannotationen auf unterschiedlichen linguistischen Ebenen, sie stellt unabhängig davon ebenfalls einen Vorteil für die Recherche in einem Lernerkorpus dar. Mit Rückgriff auf die ZH ist es möglich, im Lernerkorpus syntaktische Elemente zu ermitteln, die in der Original-Lerneräußerung fehlen, d. h. vom Lernenden ausgelassen wurden, und somit durch eine Abfrage der Original-Lerneräußerung nicht erfasst werden könnten. Ein Beispiel hierfür bietet die folgende Abbildung (Abb. 10).

O	Ich	hoffe		daine	antwort	[...]
ZH	Ich	hoffe	auf	deine	Antwort	[...]

Abbildung 10: Beispiel für die Rekonstruktion eines ausgelassenen Elements (hier: Präposition auf) auf Ebene der ZH. Entnommen aus dem Lernerkorpus MERLIN (ID: 1071_0024874). O=Original-Lerneräußerung.

Vor allem in solchen Lernerkorpora, in denen die ZH nicht nach Wortarten annotiert ist – so wie es beim Lernerkorpus MERLIN der Fall ist⁹⁵ – bietet die Verfügbarkeit von ZH einen deutlichen Mehrwert für die Korpusrecherche. So ist es möglich, bei der Korpusrecherche die Abfrageebene (O oder ZH) näher zu spezifizieren, um z. B. ausgelassene Elemente gezielt ein- oder auszuschließen.

Um mithilfe abstrakter Konzepte Abweichungen zwischen Original-Lerneräußerung und ZH ermitteln zu können, sind in manchen Lernerkorpora (z. B. in Falko und MERLIN) zusätzlich Editier-Tags angegeben. Editiertags werden automatisch erstellt und kodieren die Differenz zwischen Original-Lerneräußerung und der entsprechenden ZH (vgl. Reznicek et al. 2012: 60⁹⁶). Abbildung 11⁹⁷ illustriert die Editiertags zum Beispiel aus Abbildung 10. Die Editiertags sind auf der Annotationsebene ZHDiff repräsentiert.

⁹⁵ Das Lernerkorpus *Falko* dahingegen verfügt über wortartenannotierte ZH.

⁹⁶ In Reznicek et al. (2012: 60f.) findet man eine vollständige Liste der Editiertags, die so auch in MERLIN enthalten sind.

⁹⁷ Die Abbildungen dienen primär dazu, einzelne Editiertags zu illustrieren. Bei Angabe der Zielhypothesen wird daher auf eine nähere Spezifizierung (minimale Zielhypothese (ZH1) bzw. erweiterte Zielhypothese (ZH2)) verzichtet.

O	Ich	hoffe		daine	antwort	[...]
ZH	Ich	hoffe	auf	deine	Antwort	[...]
ZHDiff			INS	CHA	CHA	

Abbildung 11: Beispiel für die Editier-Tags *INS* (repräsentiert eine Einfügung) und *CHA* (repräsentiert eine Änderung). Entnommen aus dem Lernerkorpus MERLIN (ID: 1071_0024874).

Wie Abbildung 11 verdeutlicht, stellen die Editiertags abstrakte Repräsentationen von Abweichungen dar, und enthalten keine weitere linguistische Beschreibung der jeweiligen Abweichungen. So sind die Abweichungen von *daine* (O) zu *deine* (ZH) und *antwort* (O) zu *Antwort* (ZH) auf Ebene der Editiertags (ZH1Diff) identisch kodiert, obwohl es sich um unterschiedliche Arten von Abweichungen handelt. Im ersten Fall liegt eine Abweichung auf Graphemebene vor, im zweiten Fall handelt es sich um eine Abweichung im Bereich der Groß- und Kleinschreibung. Somit werden auf Ebene der ZHDiff Abweichungen zwischen Original-Lerneräußerung und ZH lediglich grob klassifiziert, sie bieten allerdings keine differenzierte Beschreibung der Abweichungen zwischen O- und ZH-Ebene. Bei der Korpusrecherche können die Editiertags jedoch beispielsweise mit den – wenn vorhanden – Wortartenannotationen kombiniert werden. So wäre es z. B. möglich, nach der Wortart Präposition im Lernerkorpus zu suchen, die auf Ebene der ZH verändert wurde (Editiertag CHA). Auf diese Weise ist es möglich, die abstrakten Editiertags zu konkretisieren und gezielt nach spezifischen Abweichungen zwischen Original-Lerneräußerung und ZH zu suchen, wobei auch dieser Art der Suche nach Abweichungen Grenzen gesetzt sind (vgl. auch Lüdeling/Hirschmann 2015: 146)⁹⁸.

Oben wurde erwähnt, dass mit Rückgriff auf die ZH und dem Editiertag *INS* auch solche Elemente gesucht und gefunden werden können, die in der Original-Lerneräußerung ausgelassen wurden (vgl. Abb. 11). Entsprechend dazu kann das Editiertag *DEL* genutzt werden, um nach Elementen zu suchen, die zwar in der Original-Lerneräußerung enthalten sind, auf Ebene der ZH jedoch gelöscht wurden. Abbildung 12 gibt ein Beispiel für eine hinzugefügte Präposition (hier: *bei*), die auf Ebene der ZH gelöscht wurde. Auf Ebene der ZHDiff ist die Löschung durch den Editiertag *DEL* (engl. *delete*) repräsentiert.

⁹⁸ Die Kombinierbarkeit von Editiertags und anderen Annotationsebenen (z. B. auch die Lemmatisierung) hat ihre Grenzen. Beispielsweise können spezifische orthographische Abweichungen auf diese Weise nicht ermittelt werden. Zwar ist es möglich mithilfe der Kombination aus Lemma und dem Editiertag *CHA* alle Abweichungen in der Schreibweise dieses Lemmas zu finden, jedoch werden dann z. B. Abweichungen in der Groß- und Kleinschreibung und Abweichungen in der Rechtschreibung nicht differenziert. Der Editiertag *CHA* schließt zusätzlich solche Fälle mit ein, in denen eine lexikalische Einheit (z. B. die Präposition *auf*) gänzlich ersetzt wurde (z. B. durch die Präposition *in*). Auch ist es mithilfe der Editiertags nicht möglich, gezielt nach Abweichungen in der Flexion eines Lemmas zu suchen. Die Kombination aus Lemma und dem Editiertag *CHA* schließt nämlich neben möglichen Flexionsfehlern ebenso orthographische Fehler mit ein (s. Lüdeling/Hirschmann 2015: 147).

O	ich	besuche	oft	bei	meiner	Schwägerin	[...]
ZH	ich	besuche	oft		meine	Schwägerin	[...]
ZHDiff				DEL	CHA		

Abbildung 12: Beispiel für das Editier-Tag DEL (repräsentiert eine Löschung). Entnommen aus dem Lernerkorpus MERLIN (ID: 1061_0120338).

Unabhängig davon, ob ZH explizit im Lernerkorpus angegeben werden oder nicht, bilden sie die Grundlage für eine andere Ebene der manuellen Annotation in Lernerkorpora – die Fehlerannotation.

3.3.2.2 Fehlerannotationen

Versteht man unter Annotationen „kontextualisierte Analysen“ von Daten (Lemnitzer/Zinsmeister 2015: 57), so sind Fehlerannotationen kontextualisierte Analysen von Fehlern. Unter Fehlerannotationen wird hier die Auszeichnung und Kategorisierung zielsprachlicher Abweichungen in einem Lernerkorpus verstanden. Da Annotationen immer auch Kategorisierungen sind, sind auch mit Fehlerannotationen a) eine Interpretation und b) ein Verlust an Informationen verbunden (Lüdeling/Hirschmann 2015: 136). Fehlerannotationen repräsentieren somit Interpretationen zielsprachlicher Abweichungen. Sie stellen eine Annotationsebene in einem Lernerkorpus dar, bilden jedoch auch gleichzeitig bereits grundlegende linguistische Analysen von Fehlern ab.

Fehlerannotationen zählen zu den wichtigsten Annotationen in Lernerkorpora (Díaz-Negrillo/Thompson 2013). Dies ist eng verbunden mit dem zentralen Interesse der Lernerkorpusforschung, eine umfassende Erforschung von Lernaltersprache zu ermöglichen, wozu auch die Analyse zielsprachlicher Abweichungen zählt. Dieses Interesse an Abweichungen spiegelt sich ebenfalls in den zentralen Analysemethoden der Lernerkorpusforschung wider – der computerunterstützten Fehleranalyse (Dagneaux/Deness/Granger 1998) und der kontrastiven Interlanguage-Analyse (Granger 2015). Auf beide Methoden gehe ich in Kapitel 3.4 näher ein.

Die Annotationen von Fehlern in einem Lernerkorpus setzt eine Definition und eine Identifikation von Fehlern voraus. Auf Probleme der Fehlerdefinition kann an dieser Stelle nicht im Detail eingegangen werden (s. Lennon 1991, Ellis/Barkhuizen 2005: 56ff.), diese ist jedoch eng mit der Fehleridentifikation verknüpft. Eine Identifikation von Fehlern ist nur möglich, „wenn man die Lerneräußerung mit einer angenommenen ‚korrekten‘ Äußerung vergleicht“ bzw. die Lerneräußerung im Hinblick auf die Zielsprache interpretiert (Lüdeling 2008: 126), denn „it is impossible *not* to interpret the data“ (Lüdeling/Hirschmann 2015: 141, Hervorhebung im Original). Das heißt, dass einer Fehlerannotation immer (!) eine Hypothese über die korrekte zielsprachliche Struktur zugrunde liegt. Diese Hypothese kann im Lernerkorpus in Form einer Zielhypothese (ZH, s. oben) explizit angegeben werden, bleibt jedoch in den meisten Lernerkorpora implizit bzw. kann nur auf Basis vorhandener Fehlerannotationen rekonstruiert werden (vgl. Lüdeling 2008: 126). So liegt im folgenden Beispiel aus Weinberger (2015: 54) eine Fehlerannotation ohne explizit angegebene ZH vor

(vgl. auch Maden-Weinberger 2015: 54). Die Fehlerannotation ist durch eine Markierung (Fehlertag) in spitzen Klammern repräsentiert.

die Erklärung für <MoArInGn> diese Phänomen ist einfach

Das Fehlertag *MoArInGn* markiert einen Genusfehler in der dem Fehlertag folgenden Wortform *diese*. Anhand des Fehlertags lässt sich ableiten, dass folgende ZH zugrunde liegt: *die Erklärung für dieses Phänomen ist einfach*. Ohne Angabe einer expliziten ZH muss diese a) extern geleistet werden und b) wird der Eindruck erzeugt, dass diese (implizite) ZH und damit die Kategorisierung des Fehlers als Genusfehler eindeutig ist. Dies ist jedoch nicht der Fall. Die alternative ZH *die Erklärung für diese Phänomene ist einfach* wäre hier (ohne weiteren Kontext) gleichfalls plausibel. Jedoch wäre bei dieser alternativen ZH die Abweichung in der Wortform *Phänomen* zu identifizieren und als Numerusfehler zu kategorisieren.

Dieses Beispiel illustriert, inwiefern die ZH die Fehleridentifikation und die anschließende -kategorisierung beeinflusst (s. auch die Fallstudie in Lüdeling 2008)⁹⁹. Das Beispiel zeigt zudem, dass Annotationen (hier die ZH und die darauf basierende Fehlerkategorisierung) niemals die absolute ‚Wahrheit‘ kodieren, sondern eine mögliche Interpretation der vorliegenden Daten (Lüdeling/Hirschmann 2015: 137). Aus diesem Grund ist die Angabe einer expliziten ZH für die Transparenz und Nachvollziehbarkeit von Fehlerannotationen in Lernerkorpora unabdingbar (vgl. Lüdeling 2008, Lüdeling et al. 2005). Da Fehlerannotationen kontextualisierte Analysen von Abweichungen zwischen Original-Lerneräußerung und ZH darstellen, ist ein *Fehler* somit definiert als eine Abweichung zwischen Original-Lerneräußerung und ZH. Im Korpus ist der Fehler durch einen spezifischen Fehlertag repräsentiert. Dass für eine abweichende Lerneräußerung mehrere unterschiedliche ZH angesetzt werden können, die wiederum unterschiedliche Fehler und Fehlerkategorisierungen implizieren, bedeutet jedoch nicht, Fehlerannotationen in Lernerkorpora infrage zu stellen bzw. gänzlich auf diese verzichten zu müssen. Vielmehr muss sichergestellt werden, dass die in einem Korpus vorliegenden Fehlerannotationen transparent sind und somit Analyseergebnisse, die auf Fehlerannotationen beruhen, nachvollziehbar und replizierbar sind (Lüdeling 2008: 135). Die Angabe expliziter ZH stellt dabei einen zentralen Schritt dar. Eine spezifische Korpusarchitektur, in der unterschiedliche Annotationsebenen (z. B. konkurrierende ZH und entsprechend konkurrierende Fehlerannotationen) auch repräsentiert werden können, ist dabei eine zentrale Voraussetzung (vgl. Lüdeling et al. 2005, Reznicek/Lüdeling/Hirschmann 2013)¹⁰⁰.

⁹⁹ Das beschriebene Problem tritt nicht ausschließlich in digitalen Lernerkorpora auf, sondern gilt gleichermaßen für ‚analoge‘ Fehleridentifikationen und -kategorisierungen (Lüdeling 2008: 135).

¹⁰⁰ Auf unterschiedliche Korpusarchitekturen sowie Annotationsformate (flache Annotationsformate vs. standoff-Annotationsformate) kann an dieser Stelle nicht näher eingegangen werden. Grundlegend ist der Unterschied, ob die Annotationen mit der Original-Lerneräußerung vermischt sind oder ob Original-Lerneräußerung und Annotationen in separaten Dateien abgespeichert werden. Für einen Überblick zu Annotationsformaten siehe z. B. Lüdeling et al. (2005), Lüdeling/Hirschmann (2015), Lemnitzer/Zinsmeister (2015: 99ff.), Carletta et al. (2003). Unabhängig davon, ob Fehlerannotationen in einem Lernerkorpus enthalten sind oder nicht, wird vor allem in der deutschen

Obwohl Fehlerannotationen aktuell zu den häufigsten Annotationen in Lernerkorpora zählen (van Rooy 2015: 79), hat sich bisher (noch) kein Standard für Fehlerannotationen und Fehlerannotationskategorien etabliert (vgl. Lüdeling/Hirschmann 2015, Díaz-Negrillo/Fernández-Domínguez 2006). Nach Meinung von Lüdeling/Hirschmann (2015: 155) wäre ein gewisser Grad an Standardisierung in diesem Bereich sinnvoll. Ein entsprechendes Potenzial sehen sie im Einsatz von Editier-Tags, auf die ich bereits unter 3.3.2 eingegangen bin. Die Annotation von Editier-Tags nennen Lüdeling/Hirschmann (2015: 146) *edit-distance-based tagging*. Diese Art der Fehlerannotation bietet nach Lüdeling/Hirschmann (2015) eine form-basierte Beschreibung vorliegender Fehler, ohne spezifische Fehlerkategorien abzubilden. Obwohl die form-basierte Kodierung oberflächlicher Abweichungen durch Editier-Tags ihre Vorteile hat und automatisierbar ist (vgl. Reznicek/Lüdeling/Hirschmann 2013: 115), scheint diese Art der Spezifikation allein für die Fehlerannotation nicht ausreichend (Granger 2003: 467). Die mangelnde Spezifität bei der Kodierung oberflächlicher Abweichungen wird bereits von Corder (1975: 205f.) im Kontext traditioneller Fehleranalysen kritisiert:

The traditional classification into errors of omission, addition, substitution and word order is too superficial [...] to explain difficulties. Satisfactory classifications begin with an analysis which assigns errors to levels of language description, i.e. errors of orthography or phonology, of morphology or syntax, of vocabulary, and within each level according to systems, e.g. vowel or consonant systems, tense, aspect, number, gender or case.

In der fehlenden Spezifität der Annotationen von Editiertags sehen Lüdeling/Hirschmann (2015: 155) jedoch eine Möglichkeit zur Standardisierung, wohingegen sie den Nutzen feinkörniger Fehlerkategorien für die Standardisierung bezweifeln. Den Hauptgrund für den fehlenden Nutzen standardisierter Fehlerannotationskategorien sehen Lüdeling/Hirschmann (2015) in der Abhängigkeit jener Kategorien vom spezifischen Forschungsinteresse. Díaz-Negrillo/Fernández-Domínguez (2006: 86) sehen die fehlende Standardisierung von Fehlerannotationskategorien dahingegen als Zeichen für die Notwendigkeit für „further work“ in der Lernerkorpusforschung und verbalisieren den Bedarf an einem standardisierten Fehlerannotationssystem. Überlegungen zu Kriterien eines idealen Fehlerannotationssystems finden sich z. B. in Granger (2003) und van Rooy (2015).

Wie oben erwähnt, gehören Fehlerannotationen zu den wichtigsten Annotationen in einem Lernerkorpus. Die Fehlerannotationen werden mit Hilfe einer Annotationssoftware (z. B. MMAX2, Müller/Strube 2006) i.d.R. manuell durchgeführt, was mit einem hohen Zeit- und Personalaufwand verbunden ist. Die Reliabilität von Annotationen im Allgemeinen und von Fehlerannotationen im Speziellen kann mit Rückgriff auf einen Goldstandard evaluiert werden. Ein Goldstandard repräsentiert Annotationen, die als ‚korrekt‘ gewertet und zur Evaluation neuer Annotationen herangezogen werden. Eine andere Möglichkeit der Evaluation ist der Vergleich von Annotationen, die von zwei oder mehr AnnotatorInnen angefertigt werden, und die Berechnung derer Übereinstimmung. Die Übereinstimmung wird auch Inter-Annotator-Agreement (IAA) genannt (Lüdeling/Hirschmann 2015: 148f.,

Lernerkorpusforschung für ein standoff-Annotationsformat und eine Mehr-Ebenen-Architektur für Lernerkorpora plädiert (vgl. u.a. Reznicek/Lüdeling/Hirschmann 2013).

Artstein/Poesio 2008). Ein bekanntes Maß für das IAA ist der Kappa-Koeffizient (κ), der zufallsbasierte Übereinstimmungen berücksichtigt bzw. herausrechnet (Cohen 1960, Artstein/Poesio 2008: 559ff.).

Bisher gibt es wenige Ansätze zur automatischen Fehlererkennung und -annotationen (für einen Überblick s. z. B. Leacock/Chodorow/Tetreault 2015, Leacock et al. 2014 und auch Kap. 4.2.3). Exemplarisch möchte ich hier jedoch auf Arbeiten zur automatischen Erkennung von Fehlern im Präpositionsgebrauch verweisen, die aktuell für schriftliche Produktionen von Englischlernenden vorliegen¹⁰¹. Hierzu gehören vor allem Chodorow/Tetreault/Han (2007), Tetreault/Chodorow (2008b) sowie Cahill et al. (2013). Insgesamt zeigen Arbeiten, die sich automatischen oder kostengünstigen Möglichkeiten der Fehlerannotation widmen, ein Interesse an dieser Art der Annotation und attestieren ihr somit einen hohen Stellenwert in der Lernerkorpusforschung.

3.4 Zentrale Analysemethoden der Lernerkorpusforschung

Die beiden zentralen¹⁰² Analysemethoden der Lernerkorpusforschung sind die computerunterstützte Fehleranalyse (computer-aided error analysis, CEA) und die kontrastive Interlanguage-Analyse (contrastive interlanguage analysis, CIA). Beide Methoden werde ich in diesem Kapitel näher beschreiben, da diese Methoden auch in der vorliegenden Dissertation Anwendung finden (vgl. Kap. 7.1).

3.4.1 Computerunterstützte Fehleranalyse

Die computerunterstützte Fehleranalyse hat ihre Wurzeln in der traditionellen Fehleranalyse, d. h. der Analyse von Abweichungen (=Fehlern) in lernersprachlichen Produktionen. Die Fehleranalyse hat im Bereich der Spracherwerbsforschung eine lange Tradition (vgl. Ellis/Barkhuizen 2005: 51ff.), von der auch eine Vielzahl an Veröffentlichungen zeugt (vgl. dazu Ahrenholz 2014: 170). Im Kontext der Spracherwerbsforschung war die Fehleranalyse in den 1960er Jahren dominant und erlebte ihren Höhepunkt in den 1970ern und 1980ern Jahren (vgl. Ellis 2008a: 58ff.). Als Alternative zur damals prominenten kontrastiven Analyse, die wiederum auf Annahmen der kontrastiven Hypothese¹⁰³ zum Spracherwerb beruht (vgl. Kuhberg 2001), bot die Fehleranalyse einen neuen Zugang zur Analyse von *Lernersprache* (auch Interlanguage genannt) (Selinker 1972), deren Konzept sich ebenfalls in den 70er Jahren etablierte. Die Lernersprache wird verstanden als eine sprachliche Varietät, die Lernende im

¹⁰¹ Entsprechende Ansätze für das Deutsche sind mir bisher nicht bekannt. Einen Ansatz für das Schwedische präsentieren Eeg-Olofsson/Knutsson (2003).

¹⁰² Für einen umfassenden Überblick über methodische Ansätze in der Lernerkorpusforschung siehe Callies (2015).

¹⁰³ Die kontrastive Hypothese geht davon aus, dass die Erstsprache den Erwerb einer anderen Sprache determiniert, sodass sich der Spracherwerb und insbesondere alle Fehler im Spracherwerb auf Basis des kontrastiven Vergleichs von Ausgangs- und Zielsprache vorhersagen ließen (Kuhberg 2001): Gemeinsamkeiten beider Sprachen führen zum positiven Transfer während Unterschiede zu negativem Transfer (Interferenz) führen und somit zu Fehlern in lernersprachlichen Produktionen. Diese ‚starke‘ Version der kontrastiven Hypothese wurde jedoch schnell empirisch widerlegt. So zeigten empirische Untersuchungen, dass Fehler nicht immer auf Basis des kontrastiven Vergleichs beider Sprachen eindeutig vorhergesagt werden konnten. Zudem konnte man zeigen, dass Lernenden mit verschiedenen Ausgangssprachen ähnliche Fehler unterliefen (Kuhberg 2001: 657f.).

Spracherwerbsprozess ausbilden, wobei sowohl Strukturen der Erstsprache als auch die der Zielsprache eine Rolle spielen, wie Apeltauer (2010: 833) beschreibt:

Lernersprachen sind anfangs an der Erstsprache eines Lernenden orientiert, entwickeln später aber zunehmend eigene Besonderheiten, die sich im Laufe der Zeit mehr und mehr zielsprachlichen Strukturen annähern, bis sie diese (im Idealfalle) erreichen. Unter formalen und funktionalen Gesichtspunkten kann eine Lernersprache als ein mehr oder weniger reduziertes System einer Zielsprache bezeichnet werden.

Das ‚reduzierte‘ System ist jedoch keine bloße Übertragung morphologischer und grammatischer Strukturen aus der Erstsprache unter Verwendung zielsprachlicher Vokabeln. Die Lernersprache weist ebenso systematische Abweichungen, sowohl von der Erstsprache als auch von der Zielsprache, auf (Tarone 2013: 2795).

Lernersprachen sind durch bestimmte Merkmale gekennzeichnet. Zu den Merkmalen, die eine Fehleranalyse von Lernersprache motivieren, zählt die Systematizität sowie die Variabilität. Unter das Merkmal Systematizität fällt z. B. die Beobachtung, dass Abweichungen in lernersprachlichen Produktionen systematisch sind und damit auf das Regelsystem der Lernersprache hindeuten (Apeltauer 2010: 835). Zum Merkmal Variabilität gehört z. B., dass sich die Art der Fehler im Erwerbsverlauf verändert und Fehler somit entwicklungsbedingt sind (Ellis/Barkhuizen 2005: 55). Durch eine Fehleranalyse lernersprachlicher Produktionen ist es folglich möglich, einen Einblick in das System der Lernersprache zu erhalten bzw. dieses System zu rekonstruieren sowie die lernersprachliche Entwicklung nachzuzeichnen (Corder 1975: 204f.). Fehler werden als ‚Fenster‘ in das lernersprachliche System gesehen und gelten als Hinweise auf die sprachliche Kompetenz von Lernenden (Gass/Selinker 2008: 102). Dennoch steht die Fehleranalyse allgemein in der Kritik, den Einblick in den Erwerbsprozess zu verhindern, da sie primär produkt- und nicht prozessorientiert ist (Ahrenholz 2014: 170f., Dagneaux/Denness/Granger 1998: 164). Zudem wird kritisiert, dass die Fehleranalyse ausschließlich Abweichungen fokussiert und damit diejenigen Strukturen außer Acht lässt, die von Lernenden bereits erfolgreich beherrscht werden. Ellis/Barkhuizen (2005: 70) kommen in Bezug auf die (traditionelle) Fehleranalyse zu dem Schluss, dass sie in der Spracherwerbsforschung nicht mehr die bevorzugte Methode in der Analyse von Lernersprache darstellt. Dennoch findet sie weiterhin Anwendung. Den zahlreichen Publikationen zur Fehleranalyse in den vergangenen Jahrzehnten¹⁰⁴ lässt sich entnehmen, dass die Fehleranalyse im Rahmen der Spracherwerbsforschung dennoch weiterhin eine große Rolle spielt. Dies liegt vor allem daran, dass viele WissenschaftlerInnen „in der Fehleranalyse [...] einen Zugang zur Analyse des Erwerbs von zweit- oder fremdsprachlichen Kompetenzen [sehen]“ (Ahrenholz 2014: 171).¹⁰⁵ Dies gilt vor allem für WissenschaftlerInnen, die mit Lernerkorpora arbeiten.

¹⁰⁴ Für einen Überblick s. z. B. Palmberg (1980) und Cherubim (1980), siehe auch die genannten Publikationen in Ellis/Barkhuizen (2005) und Ellis (2008: 58ff.).

¹⁰⁵ Ellis/Barkhuizen (2005: 70) erwähnen zudem, dass die Methodologie der Fehleranalyse eine direkte Relevanz für die Sprachdidaktik hat.

In der aktuellen Lernerkorpusforschung stellt die Fehleranalyse eine zentrale Analyse­methode dar (Dagneaux/Denness/Granger 1998: 173). Das Aufkommen von digitalen Lernerkorpora und die damit verbundene Anwendung korpuslinguistischer Werkzeuge und Methoden auf die Analyse von Lernerkorpusdaten verschaffte der Fehleranalyse einen neuen Auftrieb. In Form der computerunterstützten Fehleranalyse (engl. computer-aided error analysis, kurz CEA) wird sie in zahlreichen linguistischen Untersuchungen in Lernerkorpora angewandt.¹⁰⁶ Ebenfalls ist weiteres zukünftiges Interesse an der Anwendung der computerunterstützten Fehleranalyse zu erwarten¹⁰⁷.

Nach Dagneaux/Denness/Granger (1998: 165) ist die computerunterstützte Fehleranalyse eine neue Form der Fehleranalyse, bei der die Abweichungen in lernersprachlichen Äußerungen in einem Lernerkorpus identifiziert, kategorisiert (=annotiert) und interpretiert werden. Somit ist eine systematische Fehlerannotation ein integraler Bestandteil und damit ein zentraler Schritt in der computerunterstützten Fehleranalyse¹⁰⁸. Auf Fehlerannotationen in Lernerkorpora bin ich bereits unter Kapitel 3.3.2 ausführlich eingegangen. Alle dort angesprochenen Anforderungen und Herausforderungen gelten demnach ebenfalls für die Durchführung einer computerunterstützten Fehleranalyse.

Die einzelnen Schritte einer computerunterstützten Fehleranalyse sind in Dagneaux/Denness/Granger (1998) am Beispiel eines ausgewählten Lernerkorpus beschrieben und umfassen die Identifizierung, Korrektur, Kategorisierung und Annotation von Fehlern sowie die Auswertung der Fehlerannotationen in einem Lernerkorpus. Die einzelnen Schritte in der Durchführung einer computerunterstützten Fehleranalyse und die damit verbundenen Herausforderungen gleichen denen einer traditionellen Fehleranalyse (Ellis/Barkhuizen 2005: 57), jedoch unterscheiden sich beide Ansätze in einigen wesentlichen Punkten voneinander. An dieser Stelle möchte ich die Unterschiede und Gemeinsamkeiten der computerunterstützten und der traditionellen Fehleranalyse näher beleuchten. Abbildung 13 gibt einen Überblick über die Schritte beider Analyse­methoden und stellt die damit verbundenen methodologischen Überlegungen bei der traditionellen Fehleranalyse und der computerunterstützten Fehleranalyse gegenüber. Ich orientiere mich hier hauptsächlich an den Beschreibungen in Ellis/Barkhuizen (2005: 51-71) und Dagneaux/Denness/Granger (1998).

¹⁰⁶ Siehe entsprechende Publikationen in der Learner Corpus Bibliography (LCB). Dazu gehören u.a. die Arbeiten von Granger (1999), Götz (2015) und Glaznieks/Abel (2017).

¹⁰⁷ Derzeit (2018/2019) wird z. B. ein neues deutsches Lernerkorpus mit verschiedenen Fehlerannotationsebenen aufgebaut, mit dem Ziel Abweichungen in Lerneräußerungen näher zu analysieren (Hirschmann/Nolda 2019).

¹⁰⁸ Die methodologische Nähe von Fehlerannotation und Fehleranalyse zeigt sich bei Lüdeling/Hirschmann (2015). In ihrem Beitrag behandeln sie Fehlerannotationen und gehen dabei parallel u.a. ausführlich auf einzelne Schritte der computerunterstützten Fehleranalyse ein.

Schritte	Traditionelle Fehleranalyse	Computerunterstützte Fehleranalyse
Sammlung von Lerneräußerungen 1.	Informationen zur Datengrundlage oft nicht ersichtlich/nicht dokumentiert	Kontrollierte Erhebung, systematisch erstellte und dokumentierte Datensammlung (digitales, linguistisch aufbereitetes Lernerkorpus)
Fehleridentifikation 2.	<ul style="list-style-type: none"> - nur Fehler im Fokus - Identifikation abhängig von Fehlerdefinition - implizite Hypothese über zielsprachliche Struktur (Zielhypothese) - Festlegung einer Vergleichsnorm 	<ul style="list-style-type: none"> - nur Fehler im Fokus - Identifikation abhängig von Fehlerdefinition - Implizite o. explizite Hypothese über zielsprachliche Struktur (Zielhypothese) - Festlegung einer Vergleichsnorm
Fehlerbeschreibung 3.	<ul style="list-style-type: none"> Abweichung zw. impliziter Zielhypothese und Lerneräußerung - Fehlerkategorisierung, kein standardisiertes Schema 	<ul style="list-style-type: none"> Abweichung zw. expliziter o. impliziter Zielhypothese und Lerneräußerung - Fehlerkategorisierung (Fehlerannotation) mit standardisiertem Schema
Fehlererklärung 4.	<ul style="list-style-type: none"> Ermittlung der Fehlerursache (z.B. interlingual o. intralingual) - i.d.R. keine Berücksichtigung des Fehlerkontextes 	<ul style="list-style-type: none"> Ermittlung der Fehlerursache (z.B. interlingual o. intralingual) - Berücksichtigung des Fehlerkontextes - Anwendung korpuslinguistischer Methoden und statistischer Verfahren

Abbildung 13: Die traditionelle Fehleranalyse und die computerunterstützte Fehleranalyse: Gegenüberstellung.

Die einzelnen Schritte einer Fehleranalyse umfassen im Allgemeinen (1.) die Sammlung von Lerneräußerungen, (2.) die Fehleridentifikation, (3.) die Fehlerbeschreibung sowie (4.) die Fehlererklärung (s. Abb. 13)¹⁰⁹. Ich möchte mich im Folgenden auf die computerunterstützte Fehleranalyse fokussieren und die Vorteile dieses Ansatzes gegenüber dem Ansatz der traditionellen Fehleranalyse skizzieren. Ebenfalls gehe ich auf Herausforderungen der einzelnen Schritte ein, denen sich beide methodischen Ansätze stellen müssen.

Die Vorteile der computerunterstützten Fehleranalyse liegen vor allem in der kontrollierten Erhebung der zugrundeliegenden Lernerdaten in Form eines digitalen und i.d.R. linguistisch aufbereiteten Lernerkorpus (Schritt 1) sowie einem für das jeweilige Lernerkorpus standardisiertem Annotationsschema, in dem die Fehlerkategorien klar definiert sind. Die Fehlerkategorien stellen analytische Konzepte dar, die auf die im Datensatz vorliegenden Lerneräußerungen angewandt werden und eine systematische Fehlerkategorisierung bzw. -annotation gewährleisten (Tenfjord/Hagen/Johansen 2006: 102) (Schritt 3).

Zudem ermöglicht eine systematische Fehlerannotation, gezielt nach ausgewählten Fehlerkategorien im Lernerkorpus zu suchen und die annotierten Fehler in einem größeren Kontext (also nebst korrekten Lerneräußerungen) anzeigen zu lassen und auszuwerten. Im Rahmen einer computerunterstützten Fehleranalyse ist es möglich, mithilfe

¹⁰⁹ Ein weiteres Verfahren, die Fehlerbewertung, stellt keinen Schritt innerhalb einer Fehleranalyse dar, sondern dient bereits der Anwendung der Ergebnisse einer Fehleranalyse (Ellis/Barkhuizen 2005: 67). Aus diesem Grund ist die Fehlerbewertung in Abbildung 13 nicht aufgeführt.

korpuslinguistischer Analysemethoden und differenzierter statistischer Auswertungsverfahren Erklärungsansätze für frequente Fehler im Lernerkorpus zu erörtern (z. B. frequente Fehler und ihre Abhängigkeit von Lernervariablen, s. Gries 2013) (Schritt 4).

Im Rahmen der traditionellen Fehleranalyse hingegen werden zentrale Informationen zu den Lernenden und dem Erhebungskontext oft nicht (ausreichend) dokumentiert und bleiben somit oft unklar¹¹⁰ (vgl. Granger 2008: 264; Gass/Selinker 2008: 57) (Schritt 1), die Fehlerbeschreibung basiert nicht auf einem standardisiertem Schema (Schritt 3) und die einzelnen Fehler werden i.d.R. ohne größeren Kontext analysiert (Ellis/Barkhuizen 2005: 63, Granger 2008: 268) (Schritt 4).

Beide Ansätze teilen aber auch gewisse Einschränkungen und Herausforderungen in der Analyse von Lerner Sprache. Sowohl die traditionelle Fehleranalyse als auch die computerunterstützte Fehleranalyse fokussieren i.d.R. ausschließlich fehlerhafte Lerneräußerungen (Schritt 2) und verstellen damit indirekt den Blick auf die Strukturen, die Sprachlernende bereits beherrschen (Ellis/Barkhuizen 2005: 71). Anhand einer Fehleranalyse können zudem keine Aussagen gemacht werden über ‚versteckte‘ Fehler (engl. *covert errors*). Bei ‚versteckten‘ Fehlern zeigen sich an der sprachlichen Oberfläche keine Abweichungen, jedoch drückt die gewählte sprachliche Form nicht die von den Lernenden intendierte Äußerung aus. Bei einer Fehleranalyse können somit hauptsächlich nur die an der Oberfläche sichtbaren Abweichungen (engl. *overt errors*) berücksichtigt werden (Callies 2015: 41). Bei den an der Oberfläche sichtbaren Abweichungen sollte im Rahmen der Fehleranalyse zwischen Performanz- und Kompetenzfehlern unterschieden werden. Als Performanzfehler gelten z. B. Flüchtigkeitsfehler, also Fehler, die nicht aufgrund mangelnder Sprachkompetenz entstanden sind und i.d.R. retrospektiv von den entsprechenden Lernenden als solche identifiziert und korrigiert werden können. Kompetenzfehler dahingegen sind systematischer Natur, lassen Schlüsse auf zugrundeliegenden Sprachkompetenzen zu und Lernende sind sich dieser Fehler nicht bewusst (Callies 2015: 41, Ellis/Barkhuizen 2005: 62). Die Unterscheidung zwischen Performanz- und Kompetenzfehlern in lernersprachlichen Äußerungen ist jedoch nicht trivial und in einigen Fällen auch nicht eindeutig möglich. Eine direkte Einbeziehung des/der entsprechenden Lernenden – indem man überprüft, welche Fehler er/sie eigenständig korrigieren kann und welche Fehler damit Performanzfehler darstellen und welche Kompetenzfehler sind – ist diesbezüglich die beste Lösung, jedoch in der Regel nicht umsetzbar (Ellis/Barkhuizen 2005: 64). Aus diesem Grund wird oft auf eine Unterscheidung zwischen Performanz- und Kompetenzfehlern verzichtet. Bei einer computerunterstützten Fehleranalyse in einem Lernerkorpus kann i.d.R. keine Unterscheidung zwischen an der Oberfläche sichtbaren und nicht sichtbaren Fehlern sowie zwischen Performanz- und Kompetenzfehlern vorgenommen werden.

Beide Ansätze müssen mit der Herausforderung der Fehlerdefinition und Fehleridentifikation umgehen (Schritt 2). Hierbei ist zunächst zu entscheiden, *wie* ein Fehler

¹¹⁰ Nach Gass/Selinker (2008: 57) hängt dies mit der Art und Weise zusammen, wie in Publikationen über die Daten berichtet wird. Viele Studien zur traditionellen Fehleranalyse haben diesen Aspekt nicht ausreichend berücksichtigt. Die Ergebnisse solcher Studien sind somit schwer nachvollziehbar und nahezu nicht replizierbar (Ellis 2008a: 47).

definiert wird und *welche* Fehler genau analysiert werden sollen. In Bezug auf die Frage, *welche* Fehler genau analysiert werden sollen, kann sich eine Fehleranalyse z. B. grammatischen Abweichungen widmen oder stilistische Abweichungen fokussieren. Zusätzlich stellt sich für die Fehlerdefinition und Fehleridentifikation bei beiden Ansätzen die Frage, *wie* die vorgefundene Abweichung korrigiert werden kann bzw. an welcher Stelle der Fehler identifiziert wird. Dies hängt eng mit der Fehlerdefinition zusammen. Bei beiden Ansätzen gilt es, eine Vergleichsnorm festzulegen, gegen die die Lerneräußerungen evaluiert werden können (Schritt 2). Hierbei müssen u. a. das Medium (mündlich/schriftlich) sowie die zielsprachliche Varietät berücksichtigt werden (Callies 2015: 41). Unter Berücksichtigung einer Vergleichsnorm gibt es z. T. mehrere Möglichkeiten, einen Fehler in einer Lerneräußerung zu rekonstruieren (vgl. Ausführungen in Kap. 3.3). Die Rekonstruktion der Lerneräußerung, die ZH, hat einen direkten Einfluss auf die Fehleridentifikation und die daran anschließende Fehleranalyse. Im Falle einer computerunterstützten Fehleranalyse in einem Lernerkorpus gibt es jedoch – gegeben die entsprechende Korpusarchitektur – die Möglichkeit, die zugrundeliegende Zielhypothese(n) zur zielsprachlichen Struktur explizit anzugeben und damit nachvollziehbar und replizierbar zu machen (s. Ausführungen in Kap. 3.3.2).

Das Lernerkorpus MERLIN, das als Datengrundlage dieser Dissertation dient, verfügt sowohl über explizite ZH als auch über ein frei verfügbares Annotationsschema, in dem die einzelnen Fehlerkategorien beschrieben werden. Somit ist die Nachvollziehbarkeit und die Replizierbarkeit der in MERLIN durchgeführten computerunterstützten Fehleranalysen gewährleistet.

In der vorliegenden Dissertation werden entsprechend den Schwerpunkten der einzelnen Fallstudien ausgewählte Lerneräußerungen im Lernerkorpus MERLIN betrachtet (Schritt 1); Die Fehleridentifikation (Schritt 2) erfolgt auf Basis der im Lernerkorpus vorliegenden ZH; Ein Fehler wird somit definiert als eine Abweichung zwischen Original-Lerneräußerung und der vorliegenden ZH (s. Kap. 3.3.2.2). Für die Fehlerbeschreibung (Schritt 3) greife ich vorrangig auf die im Korpus vorliegenden Fehlerannotation bzw. die entsprechenden Fehlerannotationskategorien zurück. Die Erklärung der Fehler (Schritt 4) erfolgt unter Anwendung statistischer Auswertungsmethoden und unter Berücksichtigung der Leitfragen und Hypothesen, die basierend auf aktuellen Annahmen gebrauchsbasierter Ansätze zum Zweit- und Fremdspracherwerb für die jeweilige Fallstudie formuliert werden.

3.4.2 Kontrastive Interlanguage-Analyse

Die kontrastive Interlanguage-Analyse (contrastive interlanguage analysis, kurz CIA) gehört zu den am weitesten verbreiteten methodologischen Ansätzen in der Lernerkorpusforschung (Callies 2015: 39). Granger (1996) führte die kontrastive Interlanguage-Analyse als eine neue Form der kontrastiven Analyse ein, bei der muttersprachliche und lernersprachliche Produktionen ein und derselben (Ziel-)Sprache kontrastiert werden, d. h. zwei Sprachvarietäten stehen hier im Vergleich. Dies grenzt diesen methodologischen Ansatz ab von der traditionellen kontrastiven Analyse, bei der verschiedene Sprachen untersucht und miteinander kontrastiert werden (s. Lado 1957).

Die Methode der kontrastiven Interlanguage-Analyse unterschied ursprünglich zwei Arten des kontrastiven Vergleichs von Sprachvarietäten: (1) kontrastiver Vergleich von

Lernenden mit MuttersprachlerInnen und (2) kontrastiver Vergleich von Lernenden mit unterschiedlichen Erstsprachen und gleicher Zielsprache (Granger 1996: 44). In einem modifizierten Ansatz der kontrastiven Interlanguage-Analyse hat Granger (2015) die Ansätze des kontrastiven Vergleichs erweitert und einige methodologische und terminologische Änderungen vorgenommen. Der aktuelle Ansatz ist in Abbildung 14 illustriert.

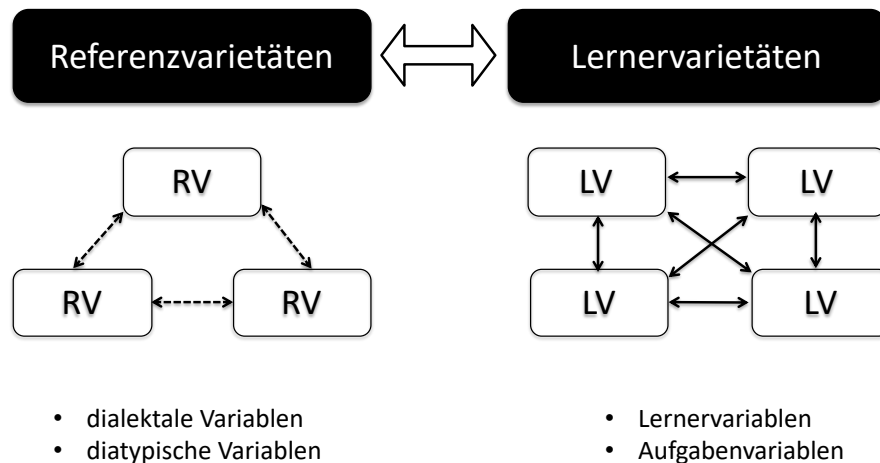


Abbildung 14: Kontrastive Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (nach Granger 2015: 17).

Im modifizierten Ansatz der kontrastiven Interlanguage-Analyse (Abb. 14) führt Granger (2015) Referenz- und Lernervarietäten (RV und LV) ein und ersetzt damit die terminologisch fehlende Präzision der ursprünglichen Begriffe *Muttersprache* und *Lernersprache*, die in ihrem ersten Ansatz (Granger 1996) prominent waren. Die Einführung des Begriffs *Referenzvarietäten*, so Granger (2015: 17), verdeutlicht die Existenz unterschiedlicher Referenzpunkte, die für einen kontrastiven Vergleich mit Lernerdaten herangezogen werden können (z. B. verschiedene Sprachvarietäten oder Kompetenzniveaus¹¹¹) (=dialektale Variablen). Zudem bezieht das Modell die Möglichkeit mit ein, simultan verschiedene Referenzvarietäten zu berücksichtigen bzw. miteinander zu vergleichen. Des Weiteren plädiert Granger (2015) in ihrem modifizierten Ansatz für eine stärkere Berücksichtigung diatypischer Variablen bei den Referenzvarietäten wie beispielweise das Medium (mündlich/schriftlich) oder die zugrundeliegende Produktionsaufgabe (s. auch Granger 2002: 12).

Der Begriff der Referenzvarietät unterstreicht zudem, dass es sich nicht zwangsläufig um eine sprachliche ‚Norm‘ handeln muss, die Lernende zu imitieren suchen wie noch im früheren Ansatz der Methode (Granger 1996) suggeriert wird. Die terminologische Umbenennung von *Lernersprache* zu *Lernervarietäten* stellt Granger (2015) nach eine Aufforderung dar, die Variabilität von Lernersprache(n) in zukünftigen kontrastiven Analysen stärker zu berücksichtigen („highlight the highly variable nature of interlanguage“, 19). Bei der Untersuchung von Lernervarietäten im Rahmen einer kontrastiven Interlanguage-Analyse sind zudem unterschiedliche Lerner- und Aufgabenvariablen zu beachten (Granger 2015: 18) (s. Abb. 14). Hierzu gehört beispielweise die Erstsprache, aber auch das Kompetenzniveau der

¹¹¹ Granger (2015: 17) schlägt z. B. vor, Lernerproduktionen anstatt mit muttersprachlichen Produktionen mit Produktionen von sprachlich kompetenten Lernenden zu vergleichen.

Lernenden sowie die zugrundeliegende Produktionsausgabe oder der Erhebungskontext. Meist liegen die entsprechenden Informationen in den Metadaten des entsprechenden Lernerkorpus vor und können somit im Rahmen einer kontrastiven Interlanguage-Analyse entsprechend berücksichtigt werden.

Bei der kontrastiven Interlanguage-Analyse können a) quantitative Unterschiede im Sinne eines Mehr- und Mindergebrauchs (engl. over-/underuse) sprachlicher Phänomene und/oder b) qualitative Unterschiede im Sinne von Abweichungen (Fehlern) (engl. misuse) fokussiert werden (Callies 2015: 40). Dabei werden genau genommen die

Frequenzen einer gegebenen Kategorie (Wörter, Wortarten, Wortartenfolgen, Fehler einer bestimmten Art, etc.) in einem Lernerkorpus mit Frequenzen derselben Kategorie in einem anderen Korpus [...] verglichen (Lüdeling/Walter 2010: 319).

Der Vergleich kann sich dabei auf eine Gegenüberstellung einer Referenzvarietät mit einer Lernervarietät beziehen (RV vs. LV) oder auf eine Gegenüberstellung von unterschiedlichen Lernervarietäten (LV vs. LV). Als zu vergleichende Kategorien kommen, wie Lüdeling/Walter (2010: 319) erwähnen, unterschiedliche sprachliche Ebenen in Frage, wobei für die Analyse gewisser Kategorien (wie z. B. Wortarten oder bestimmte Fehlertypen) annotierte Korpora notwendig sind. Granger (2008: 267) erwähnt, dass die meisten kontrastiven Interlanguage-Analysen auf nicht-annotierten (Lerner-)Daten basieren. Mitunter könnte dies auch der Grund dafür sein, dass die kontrastive Interlanguage-Analyse zu den am weitesten verbreiteten methodologischen Ansätzen in der Lernerkorpusforschung zählt. Anders als die computerunterstützte Fehleranalyse (s.o.), deren inhärenter Bestandteil die Fehlerannotation ist, kann die kontrastive Interlanguage-Analyse bereits auf Basis einzelner Wortformen durchgeführt werden. Da der Rückgriff auf Annotationen (z. B. Wortartenannotationen) jedoch umfassendere Untersuchungsmöglichkeiten in Korpora bietet, wird auch für kontrastive Interlanguage-Analysen in letzter Zeit vermehrt auf Korpusannotationen zurückgegriffen (vgl. z. B. Zeldes/Lüdeling/Hirschmann 2008, Hirschmann et al. 2013, Breckle 2014). Exemplarisch seien hier auch Hirschmann et al. (2013) genannt, die die Vorteile syntaktisch annotierter Daten für die kontrastive Interlanguage-Analyse zum Gebrauch von Modifikatoren durch MuttersprachlerInnen und fortgeschrittene DaF-Lernende zeigen.

Der Vergleich bestimmter Kategorien in einem Lernerkorpus mit Daten eines Referenzkorpus (RV vs. LV) zählt zu den häufigsten Ansätzen innerhalb der Methodologie der kontrastiven Interlanguage-Analyse (vgl. Granger 2002: 18). Solch ein Vergleich dient dazu, Merkmale zu extrahieren, in denen sich beide Varietäten unterscheiden bzw. die charakteristisch für die untersuchte Lernervarietät sind. Durch einen Vergleich von Lernervarietäten (LV vs. LV) kann ermittelt werden, ob gewisse lernersprachliche Merkmale von spezifischen Lernervariablen abhängen, d. h. lediglich für z. B. eine Lernergruppe, einen Gebrauchskontext oder ein Kompetenzniveau charakteristisch sind (und damit z. B. auf einen Einfluss dieser Variablen hindeuten) oder ob die Ergebnisse unabhängig von einzelnen Lernervariablen und damit verallgemeinerbar sind (Callies 2015: 40). Somit dient diese Art des Vergleichs (LV vs. LV) einem besseren Verständnis von Lerner Sprache und den Einflussfaktoren im Spracherwerb (vgl. Granger 2002: 13, Granger 2009: 18). Oft werden auch

beide Ansätze kombiniert: In solchen Fällen erfolgt zunächst ein Vergleich von Lernervarietäten mit Referenzvarietyäten (hier: Daten von MuttersprachlerInnen), um auffällige Strukturen innerhalb der Lernervarietäten zu ermitteln. Mit auffälligen Strukturen sind hier solche gemeint, für die sich ein Mehr- oder Mindergebrauch bei Lernenden feststellen lässt. Sind diese Strukturen ermittelt, können sie innerhalb unterschiedlicher Lernervarietäten, und ohne weiteren Einbezug einer Referenzvarietyät, näher untersucht werden (Granger 2009: 18).

Beide Ansätze (RV vs. LV und LV vs. LV) sind mit methodischen Herausforderungen verbunden. Diese sind zum einem die Vergleichbarkeit der einzelnen Korpora und zum anderen die grundsätzliche Kritik am Vergleich einer Lernervarietyät mit einer muttersprachlichen Referenzvarietyät. Die Herausforderung der Korpusvergleichbarkeit wird von Callies (2015: 40) als „serious but yet unresolved issue“ beschrieben und betrifft beide Ansätze innerhalb der Methodologie (RV vs. LV und LV vs. LV). Da im Rahmen der kontrastiven Interlanguage-Analyse verschiedene Korpora miteinander verglichen werden, sollten die „Vergleichskorpora in möglichst vielen Designparametern übereinstimmen“ und sich im Idealfall nur in dem zu untersuchenden Parameter unterscheiden (Lüdeling/Walter 2010: 319). Im Falle des Ansatzes RV vs. LV stellt sich hier zum Beispiel die Frage, welche Referenzvarietyät sich für einen Vergleich eignet, ob man z. B. Daten von MuttersprachlerInnen oder Daten von Lernenden mit einem hohen Kompetenzniveau als Maßstab wählt. Die Wahl muttersprachlicher Daten wirft die zusätzliche Frage auf, welche Sprachvarietyät im Einzelnen herangezogen wird (Callies 2015: 40). Mit der Frage der Korpusvergleichbarkeit ist somit auch die Frage der Sprachnorm bzw. Sprachvarietyät verbunden, gegen die die Lernervarietyät evaluiert wird. Ich gehe unten näher auf diesen Punkt ein. Die Wahl des Vergleichskorpus hat somit einen direkten Einfluss auf die Datenanalyse und die Interpretation der Ergebnisse. Beim Ansatz LV vs. LV spielt die Frage nach der Korpusvergleichbarkeit insofern eine Rolle, als auch bei diesem Ansatz gewisse Parameter vergleichbar sein müssen und die zu untersuchenden Parameter (z. B. Kompetenzniveau) unterschiedlich sein sollten. Callies (2015: 52) macht z. B. darauf aufmerksam, dass insbesondere der Parameter Kompetenzniveau häufig nicht ausreichend differenziert ist bzw. meist nach lernerzentrierten Kriterien (wie z. B. Studienjahre) festgelegt wird, die die Sprachkompetenz nur unzureichend widerspiegeln (vgl. Carlsen 2012: 166). Dies hat zur Folge, dass Lernervarietyäten unter der Annahme miteinander verglichen werden, unterschiedliche, klar voneinander abzugrenzende Kompetenzniveaus zu repräsentieren, die Lernervarietyäten jedoch ggf. jeweils individuelle Varianzen hinsichtlich des Kompetenzniveaus aufweisen und sich daher nicht für einen adäquaten kontrastiven Vergleich eignen. Zusätzlich ist für beide Ansätze (RV vs. LV und LV vs. LV) zu beachten, dass die Vergleichbarkeit andere Parameter, wie z. B. die zugrundeliegende Textproduktionsaufgabe bzw. Aufgabenstellung ebenfalls berücksichtigt werden muss.

Das Problem der Korpusvergleichbarkeit beim Ansatz RV vs. LV ist eng mit Frage verbunden, gegen welche Sprachnorm eine Lernervarietyät evaluiert wird. Der Ansatz, Produktionen von MuttersprachlerInnen mit Produktionen von Lernenden zu vergleichen, wird oft kritisiert (s. dazu z. B. Ausführungen in Granger 2015). Die Hauptkritik betrifft dabei

den Vergleich selbst bzw. die impliziten Annahmen, die diesem Vergleich zugrunde liegen. Wie eingangs erwähnt, stellt die Lernervarietät – die Lernaltersprache – ein eigenständiges Sprachsystem dar, das sich systematisch sowohl von den Strukturen der Erstsprache als auch von den Strukturen der Zielsprache unterscheidet. Durch einen Vergleich von lernaltersprachlichen Produktionen mit denen von MuttersprachlerInnen wird die eigenständige Sprachvariante der Lernenden nach z. B. Bley-Vroman (1983), Selinker (2014: 230) und Larsen-Freeman (2014: 217) nicht ausreichend anerkannt, sondern (lediglich) als eine unvollkommene Version der Zielsprache angesehen, der sie in einem Vergleich gegenübergestellt wird. Bley-Vroman (1983) äußert sich kritisch gegenüber jeglicher Art¹¹² des Ver- und Abgleichs der sprachlichen Produktionen von Sprachlernenden und MuttersprachlerInnen. Für die Problematik dieses Vergleichs prägte Bley-Vroman (1982: 2) den Begriff der ‚comparative fallacy‘. In Arbeiten zur kontrastiven Interlanguage-Analyse wird dieser Begriff und die dahinterstehende Kritik bewusst aufgegriffen und reflektiert. Granger (2015) begegnet der Kritik der ‚comparative fallacy‘, indem sie u. a. mit dem neuen Ansatz zur kontrastiven Interlanguage-Analyse (s. Abb. 14) und den darin enthaltenen terminologischen Anpassungen deutlich macht, dass die muttersprachliche Referenzvariante *eine*, jedoch nicht die einzig mögliche Vergleichsbasis darstellt.

In der Lernaltersprachforschung wird der Vergleich von Daten Sprachlernender mit denen von MuttersprachlerInnen dennoch wiederkehrend verteidigt. Granger (2015: 13) macht beispielsweise deutlich, dass der Vergleich von Lernenden mit MuttersprachlerInnen in vielen empirischen Untersuchungen im Kontext der Spracherwerbsforschung Anwendung findet, ohne aber explizit thematisiert zu werden. So weist Granger (2009: 18) z. B. darauf hin, dass Vergleiche von Lernenden mit unterschiedlichen Kompetenzniveaus ebenfalls auf Annahmen zu einer (muttersprachlichen) Zielnorm beruhen, auf deren Basis die Kompetenzniveaus erst ihre inhaltliche Aussagekraft erhalten. Ähnliches gilt für Untersuchungen, in denen Grammatikalitätsurteile eine Rolle spielen, so Granger (2015: 14): Auch hier werden (mutter)sprachliche Zielnormen präsupponiert. Granger (2009: 19) erwähnt weiterhin, dass in der traditionellen Spracherwerbsforschung (die i. d. R. nicht auf Daten aus digitalen Lernaltersprachkorpora zurückgreift) die muttersprachliche Zielnorm meist implizit bleibt bzw. auf Intuitionen beruht, während sie in der Lernaltersprachforschung zumindest explizit und korpusbasiert, und zwar durch die Wahl des Vergleichskorpus, bestimmt wird und damit zumindest transparent und im besten Fall auch replizierbar ist.

Die Referenz auf eine muttersprachliche (Ziel-)Norm spielt auch eine wichtige Rolle im Kontext des Fremdsprachenunterrichts: Granger (2002: 13) verteidigt den Vergleich von MuttersprachlerInnen und Sprachlernenden mit dem Verweis, dass das Ziel eines Fremdsprachenunterrichts darin besteht, die sprachliche Kompetenz von Lernenden zu verbessern, und dies wiederum bedeutet, die Kompetenzen der Lernenden an die von MuttersprachlerInnen anzunähern. Somit spiegelt der Vergleich von Lernenden mit MuttersprachlerInnen die Lebensrealität von Sprachlernenden wider und kann pädagogisch

¹¹² Hierzu zählt auch die (computerunterstützte) Fehleranalyse, bei der die Fehler auf Basis einer zielsprachlichen Norm identifiziert und interpretiert werden.

wertvoll sein (Granger 2009: 19). Eine ähnliche Beobachtung liefern auch Ellis/Barkhuizen (2005: 360): Weil die Leistungen von Sprachlernenden typischerweise auf Basis zielsprachlicher Normen evaluiert werden, führt dies bei den Sprachlernenden dazu, dass diese eigenständig kognitive Vergleiche („*cognitive comparisons*“) zwischen ihren Sprachproduktionen und den zielsprachlichen Normen durchführen. Ellis/Barkhuizen (2005: 360) schließen daraus, dass der Vergleich von Lernenden mit MuttersprachlerInnen somit auch psycholinguistisch valide ist. Weiterhin machen sie jedoch darauf aufmerksam, dass bei einem solchen Vergleich Daten aus ähnlichen Aufgabenstellungen herangezogen werden müssen, denn eine ‚muttersprachliche Perfektion‘ im Sprachgebrauch sollte nicht einfach implizit angenommen werden (Ellis/Barkhuizen 2005: 362).

Abschließend lässt sich festhalten, dass bisher noch keine überzeugenden Argumente vorliegen, die gegen einen Vergleich von Lernenden mit MuttersprachlerInnen in der Lernerkorpusforschung sprechen (Granger 2015: 14). Diese Beobachtung wird auch von zahlreichen Publikationen in der Lernerkorpusforschung gestützt, die die kontrastive Interlanguage-Analyse anwenden (u.a. Breckle 2014, Breindl 2016, Zeldes/Lüdeling/Hirschmann 2008, Gilquin/Paquot 2008).

In der vorliegenden Dissertation steht der Vergleich von Lernervarietäten (LV vs. LV) im Vordergrund. Da jedoch hauptsächlich die Ermittlung qualitativer Unterschiede (Fehler) unter Berücksichtigung unterschiedlicher Sprachkompetenzniveaus im Vordergrund steht, liegt ebenfalls implizit ein Vergleich einer Lernervarietät mit einer Referenznorm (LV vs. RV) vor. Die Referenznorm ist durch die im Lernerkorpus vorliegenden ZH repräsentiert.

3.5 Schnittstellen der Lernerkorpusforschung

Die Lernerkorpusforschung hat unterschiedliche Anwendungsbereiche und verfügt damit über Schnittstellen zu anderen Forschungsfeldern (vgl. Abb. 5). Zu den Schnittstellen gehören die Korpuslinguistik, der Zweit-/Fremdsprachenunterricht, die Zweit-/Fremdspracherwerbsforschung sowie die maschinelle Sprachverarbeitung (natural language processing, NLP) (Granger et al. 2015: 3). Im Folgenden möchte ich auf die einzelnen Schnittstellen eingehen und auf weiterführende Literatur verweisen, wobei ein Fokus auf der Schnittstelle zur Zweit-/Fremdspracherwerbsforschung liegt.

3.5.1 Korpuslinguistik

Die Korpuslinguistik kann als eine besondere Art der Schnittstelle zur Lernerkorpusforschung betrachtet werden, da sich die Lernerkorpusforschung in den 1980er Jahren als Zweig der Korpuslinguistik entwickelt hat (Granger et al 2015: 1). Somit sind die Korpuslinguistik und die Lernerkorpusforschung sehr eng miteinander verbunden und die Lernerkorpusforschung wurde (bzw. wird) maßgeblich von KorpuslinguistInnen mit entsprechender Expertise geprägt. Dies gilt sowohl in Bezug auf den Aufbau von Lernerkorpora als auch hinsichtlich der Aufbereitung der Lernerkorpusdaten durch korpuslinguistische Werkzeuge, der Extraktion der Daten und deren (quantitativer) Auswertung (Granger 2012: 8, Granger 2008: 265ff., vgl. auch Lemnitzer/Zinsmeister 2015: 90ff.). Es ist in Zukunft zu erwarten, dass die Korpuslinguistik die Lernerkorpusforschung weiterhin beeinflussen wird. So gibt es in der Korpuslinguistik Initiativen zur Etablierung einheitlicher Repräsentationsformate für Korpusdaten. Eine solche

Initiative ist die Text Encoding Initiative (TEI)¹¹³, die seit den frühen 1990er Jahren nach der Entwicklung eines einheitlichen Standards zur Kodierung maschinenlesbarer Daten strebt. Die TEI Standards wurden bisher kaum für die Repräsentation von Lernerkorpusdaten beachtet, erste Ansätze hierzu können jedoch nun beobachtet werden¹¹⁴. Weiterhin wird auch das Potenzial der Standardisierung von Metadaten oder Annotationen in Lernerkorpora bereits diskutiert (vgl. Gilquin 2015: 31).

3.5.2 Maschinelle Sprachverarbeitung

Nach Meurers (2015: 537) überschneiden sich die Lernerkorpusforschung und die maschinelle Sprachverarbeitung (natural language processing, kurz NLP) in der Repräsentation und automatischen Analyse von Lernerksprache. Es lassen sich dabei nach Meurers (2015: 537f.) drei Nutzungsszenarien maschineller Sprachverarbeitung ausmachen, die Lernerkorpora involvieren: 1) NLP-Werkzeuge (z. B. automatische Annotationswerkzeuge) werden zur Aufbereitung der Daten in einem Lernerkorpus eingesetzt, 2) NLP-Werkzeuge werden für spezifische Analysen von Lernerkorpusdaten eingesetzt (z. B. zur automatischen Ermittlung des Kompetenzniveaus) und 3) ein Lernerkorpus wird genutzt, um NLP-Werkzeuge zu trainieren, um diese z. B. im Kontext einer Sprachlernsoftware (engl. computer-assisted language learning, CALL) zu implementieren (s. auch Kap. 3.5.3).

Die Anwendung automatischer NLP-Werkzeuge zur Aufbereitung von (teils fehlerhaften) Lernerkorpusdaten ist fehleranfällig, da NLP-Werkzeuge zur Annotation der jeweiligen (i.d.R. fehlerfreien) Zielsprache entwickelt wurden (vgl. Meurers 2015: 538). Eine mögliche Lösung hierfür ist die Entwicklung spezieller Annotationswerkzeuge für Lernerkorpora. Erste Ansätze hierzu habe ich in Kapitel 3.3.1 aufgeführt.

NLP-Werkzeuge können eingesetzt werden zur automatischen Analyse von Lernerkorpusdaten. Hancke/Meurers (2013) präsentieren einen Ansatz zur automatischen Klassifizierung von Lernertexten hinsichtlich der repräsentierten Kompetenzniveaus. Derartige Klassifizierungen können wiederum z. B. für automatische Textbewertungs- und -benotungsprogramme von Nutzen sein. Einen Überblick hierzu bieten z. B. Higgins/Ramineni/Zechner (2015). Sie erläutern den Einsatz von Lernerkorpora für die Entwicklung von Werkzeugen zur automatischen Bewertung (i.S.v. Punktevergabe) von Antworten in offenen Aufgabenformaten im Kontext von Sprachtests (mündlich/schriftlich).

Der Einsatz von Lernerkorpusdaten als Trainingsdaten für NLP-Werkzeuge ist vor allem im Kontext automatischer Fehlererkennungssoftware relevant. Lernerkorpusdaten können entweder als Trainingsdaten für die Entwicklung entsprechender NLP-Werkzeuge dienen oder zur Evaluierung dieser Werkzeuge eingesetzt werden (Gamon et al. 2013). Auf Basis von Lernerkorpora können zudem häufige Fehler von Sprachlernenden identifiziert werden, die eine automatische Fehlererkennungssoftware erkennen sollte. Die identifizierten Fehler können die Entwicklung entsprechender Werkzeuge zusätzlich leiten. Ansätze zur Entwicklung von NLP-Werkzeugen zur automatischen Erkennung ausgewählter Fehler im

¹¹³ <http://www.tei-c.org/> (Mai 2019).

¹¹⁴ Gilquin (2015: 31) erwähnt ein Lernerkorpus mit Englisch als Zielsprache, in dem die Annotationen nach TEI-Richtlinien kodiert sind.

Präpositionsgebrauch von Lernenden mit Englisch als Zielsprache werden in Kapitel 4.2.3 vorgestellt.

NLP-Werkzeuge zur automatischen Fehlererkennung können in eine Sprachlernsoftware implementiert werden (vgl. Granger 2003). Für solch einen Einsatz muss das entsprechende Werkzeug an die Daten angepasst werden. Der Einsatz von Lernerkorpora bzw. die Auswertung lernersprachlicher Daten ist hier unabdingbar. Werkzeuge zur automatischen Fehlererkennung, die in einer Sprachlernsoftware integriert werden sollen, gilt es auf spezifische, lernersprachliche Fehler zuzuschneiden (Leacock/Chodorow/Tetreault 2015: 567). Leacock et al. (2014) bieten fürs Englische einen umfassenden Überblick zu Ansätzen der automatischen Fehlererkennung für Lernende in den Bereichen Grammatik und Orthographie.

3.5.3 Zweit-/Fremdsprachenunterricht

Die Schnittstelle von Lernerkorpusforschung und Zweit-/Fremdsprachenunterricht meint den Einsatz von Lernerkorpora bzw. die Auswertung von Lernerkorpusdaten im Kontext der Sprachvermittlung. Hierbei kann man zwischen dem direkten und indirekten Einsatz von Lernerkorpora in der Sprachvermittlung unterscheiden. Direkter Einsatz bedeutet hier, dass Lernende im Sprachunterricht aktiv mit Daten aus einem Lernerkorpus arbeiten, indirekter Einsatz meint dahingehen z. B. die Entwicklung bzw. Modifizierung von Lern- und Lehrmaterialien auf Basis der Auswertung von Lernerkorpusdaten (Granger 2015: 486). Der Fokus dieses Abschnitts liegt auf dem indirekten Einsatz von Lernerkorpora, für einen Überblick zum direkten Einsatz von Lernerkorpusdaten im Sprachunterricht siehe z. B. Gilquin/Granger (2010) und Chambers (2015). Lüdeling/Walter (2010: 318) weisen darauf hin, dass „[d]ie Nutzung von Lernerkorpora in der Vermittlung, bspw. zur Bewusstmachung von Lernschwierigkeiten oder zur Erstellung von Lehrmaterialien, noch ganz am Anfang [steht]“; es lassen sich jedoch bereits eine Vielzahl an Publikationen in diesem Bereich ausmachen (siehe Granger (2008: 270ff.) für einen ersten Überblick). Für Deutsch als Zielsprache zeigen z. B. Krummes/Ensslin (2012) wie die Auswertung von Lernerkorpusdaten gewinnbringend für den DaF-Unterricht eingesetzt werden kann¹¹⁵.

Beim indirekten pädagogischen Einsatz von Lernerkorpora kann man nach Granger (2015: 488) zwei zentrale Einsatzbereiche unterscheiden: (a) die Erstellung von Referenz- und Nachschlagewerken für Lernende (Wörterbücher, Grammatiken) und (b) die Konzipierung von Lehrmaterialien (Lehrbücher und Lern-Software). Einen guten Überblick zu Einsatzbereichen (a) und (b) bietet Granger (2015, 2009: 19ff., 2003). Zusätzlich zähle ich an dieser Stelle auch (c) den Bereich der Sprachtestforschung zum pädagogischen Einsatz von Lernerkorpora.

Die Auswertung von Lernerkorpusdaten kann typische Fehler oder stilistische Abweichungen im Gebrauch der Zielsprache durch Lernende aufdecken und auf diese Weise die Erstellung von Referenz- und Nachschlagewerken wie z. B. Wörterbüchern und Grammatiken informieren. Rundell/Granger (2007) präsentieren solch einen Ansatz für die zweite Auflage des *Macmillan English Dictionary for Advanced Learners*. Hier leitete die

¹¹⁵ Zum Einsatz von muttersprachlichen Referenzkorpora in der Vermittlung von Deutsch als Fremdsprache vgl. vor allem Fandrych/Tschirner (2007), Schmidt (2008) und Lüdeling/Walter (2009).

Auswertung von Lernerkorpusdaten die Auswahl und Konzipierung von ‚Hinweis-Boxen‘ in einem Lerner-Wörterbuch, beispielsweise zu fehleranfälligen Valenzstrukturen ausgewählter Verben oder zu frequenten Kollokationen (Rundell/Granger 2007: 18). Weitere Informationen zu diesem Projekt bieten Gilquin/Granger/Paquot (2007). Lernerkorpora können ebenfalls einen Einfluss auf die Konzipierung von Lernergrammatiken haben. Einige (englischsprachige) Beispiele finden sich bei Granger (2015: 491f.). Auch hier manifestiert sich der Einfluss von Lernerkorpora bzw. deren Auswertung größtenteils in ‚Hinweis-Boxen‘ zu fehleranfälligen grammatischen Strukturen; Granger (2015: 493) nennt jedoch auch neue Ansätze in der lernerkorpus-informierten Konzipierung von Lernergrammatiken, die über Abweichungen hinausgehen und sich stattdessen auf die Sequenzierung grammatischer Inhalte im Sprachunterricht fokussieren. Ein Beispiel für solch einen Ansatz ist das EU-finanzierte Forschungsprojekt *English Grammar Profile*¹¹⁶ (Harrison 2015).

Die Auswertung fehlerannotierter Lernerkorpusdaten kann weiterhin ertragreich sein für die Konzipierung und Anpassung einer Sprachlernsoftware (computer-assisted language learning, CALL). Ein Beispiel hierfür bietet Granger (2003). Sie beschreibt ein Projekt (FreeText Project), in dem die Auswertung eines fehlerannotierten Lernerkorpus nicht nur die Erstellung und Auswahl von Übungsaufgaben in einer Sprachlernsoftware maßgeblich lenkte, sondern auch die Konzipierung eines automatischen Feedbacksystems zur Erkennung der häufigsten Fehler (Granger 2003: 475):

Within the framework of the FreeText project, error statistics and analyses were used to target the CALL exercises on learner-attested difficulties and to develop a learner corpus-informed error diagnostic tool capable of providing learners with automatic feedback on some of their most frequent errors.

Einen Überblick über weitere Sprachlernsoftware-Projekte bietet z. B. Granger (2008: 271). Die Entwicklung automatischer Fehlererkennungs- und -korrektursysteme, wie sie oben von Granger (2003: 475) angesprochen wird, steht an der Schnittstelle zum Anwendungsbereich der maschinellen Sprachverarbeitung (s. Kap. 3.5.2).

Eine weitere pädagogischen Einsatzmöglichkeit von Lernerkorpora ist der Bereich der Sprachtestforschung. Barker/Salamoura/Saville (2015) bieten einen umfassenden Überblick zur internationalen Forschung in diesem Bereich, wobei sie insbesondere auf die unterschiedlichen Ansätze eingehen, wie Lernerkorpora die Entwicklung von Sprachtests beeinflussen können. So erwähnen Barker/Salamoura/Saville (2015: 521) beispielweise, dass die Auswertung von Lernerkorpusdaten die spezifischen sprachlichen Kompetenzen von Lernenden auf bestimmten Kompetenzniveaus illustrieren kann. Erkenntnisse hierzu können genutzt werden, um über die Inhalte entsprechender Sprachtests zu entscheiden. Weiterhin können Lernerkorpora bzw. die Analyse lernersprachlicher Produktionen u.a. zur Validierung von in Sprachtests angesetzten Kompetenzniveaus eingesetzt werden (s. Wisniewski 2017 zur Diskussion des Potenzials von Lernerkorpora für den Sprachtestbereich).

¹¹⁶ Siehe auch <http://www.englishprofile.org/> (Mai 2019).

3.5.4 Zweit-/Fremdspracherwerbsforschung

Die Schnittstelle zur Zweit- und Fremdspracherwerbsforschung war und ist neben der Sprachvermittlung eine der zentralen Schnittstellen der Lernerkorpusforschung. Für die Entstehung der Lernerkorpusforschung spielte der Bezug zur Zweit- und Fremdspracherwerbsforschung eine zentrale Rolle, wie Granger (2012: 7) beschreibt:

The idea of compiling learner corpora – computerized databases of foreign or second learner language - and applying corpus linguistic tools and methods to analyze them arose from the wish to bring to the field of second language acquisition (SLA) the same kinds of benefits that corpora were providing to the linguistic field.

Die Erstellung von Lernerkorpora und deren Auswertung mittels korpuslinguistischer Methoden war von Beginn an geleitet vom Anspruch, eine umfassende Beschreibung von Lernaltersprache zu ermöglichen (Granger 1993: 57, Gilquin 2015: 9) und Lernerkorpora somit im Kontext der Zweit- und Fremdspracherwerbsforschung (Z/FSE-Forschung) einzusetzen.

Der Einsatz von Lernerkorpora im Kontext der Z/FSE-Forschung ist motiviert durch ähnliche Zielsetzungen beider Forschungsfelder. Das Erkenntnisinteresse der Z/FSE-Forschung ist ein besseres Verständnis des Wesens der Lernaltersprache, ihrer Entwicklung und der Faktoren, die beides beeinflussen (Myles 2015: 310, Hufeisen/Riemer 2010: 738). Dieses Erkenntnisinteresse entspricht einem zentralen Anliegen der Lernerkorpusforschung, nämlich „to describe and explain the nature and development of interlanguages“ (Callies 2015: 50) bzw. „[to] gain[-] a better understanding of the mechanisms of foreign or second language acquisition“ (Granger et al. 2015: 3), wobei die Lernerkorpusforschung im Vergleich zur traditionellen¹¹⁷ Z/FSE-Forschung einen stärkeren Anwendungsbezug aufweist (Granger et al. 2015: 3). Für beide Forschungsfelder sind zudem „good learner data“ (Myles 2015: 309), d. h. Daten von Sprachlernenden, essenziell.

Im Kontext der Z/FSE-Forschung repräsentieren die Daten in einem Lernerkorpus einen bestimmten Datentyp, den Granger (2012: 8) zwischen natürlichen Sprachdaten und klinischen Daten verortet. Lernerkorpora enthalten in der Regel schriftliche Produktionen, die auf Basis offener Aufgabenstellungen elizitiert wurden. Diese Art der Produktionen repräsentiert zwar keinen natürlichen Sprachgebrauch i.e.S – da die Sprachdaten nicht in einem natürlichen, kommunikativen Umfeld produziert werden – sie können u.U. dennoch als authentische Sprachdaten gelten, denn

Authentizität meint in der Lernerkorpuslinguistik, dass die Lerneräußerungen nicht allein für eine linguistische Untersuchung angefertigt wurden, sondern in einem natürlichen Kontext, und dies heißt, einem kommunikativen Kontext (Walter/Grommes 2008: 16).

Der ‚natürliche Kontext‘ ist hier zu verstehen als repräsentativ für Unterrichtsaktivitäten und daher ‚natürlich‘ für die Sprachlernsituation (Granger 2002: 8). So kann solch ein Kontext z. B. auch ein standardisierter Sprachtest sein (Walter/Grommes 2008: 15). Lernerkorpora

¹¹⁷ Unter ‚traditionell‘ verstehe ich hier die Z/FSE-Forschung, die nicht auf digitale und linguistisch aufbereitete Lernerkorpora zurückgreift. Auch wenn Datensammlungen aus der traditionellen Z/FSE-Forschung heutzutage z. T. in Form digitaler Korpora vorliegen, so enthalten diese Korpora i.d.R. keine strukturierten Daten und repräsentieren nur wenige Lernende (vgl. Skiba 2008).

enthalten Produktionsdaten von Lernenden und spiegeln den Sprachgebrauch dieser wider. Abhängig von der zugrundeliegenden Spracherwerbstheorie und den Annahmen zum Verhältnis von Sprachgebrauch und Sprachwissen (vgl. Goschler/Stefanowitsch 2014: 351) können Lernerkorpusdaten als Datenbasis in der Z/FSE-Forschung bereichernd sein oder gänzlich abgelehnt werden. Goschler/Stefanowitsch (2014: 354) weisen beispielsweise darauf hin, dass der Einsatz von Lernerkorpora nur dann sinnvoll ist, wenn zwischen Sprachgebrauch und Sprachwissen nicht unterschieden wird. In Bezug auf lernersprachliche Untersuchungen im Bereich der Grammatik merkt Rankin (2015: 233) an, dass Lernerkorpusdaten nicht gut mit Spracherwerbstheorien der generativen Grammatik¹¹⁸ vereinbar sind, jedoch im Kontext kognitiver Spracherwerbstheorien eingesetzt werden können. Myles (2015: 312) erwähnt zudem, dass Lernerkorpora als Datengrundlage besser geeignet sind für empirische Untersuchungen zu Merkmalen des linguistischen Systems von Lernenden (=der Lernaltersprache) sowie zu inter- und intralingualen Einflüssen im Sprachgebrauch. Lernerkorpora eignen sich jedoch weniger gut für Untersuchungen zu individuellen Differenzen (Motivation, Sprachlerneignung etc.) oder dem Einfluss des sozialen Umfelds auf den Sprachgebrauch und die Sprachentwicklung.

In Kontexten, in denen Lernerkorpora angemessen für Fragestellungen der Z/FSE-Forschung eingesetzt werden können, bieten sie ein großes Potenzial (Granger 2012: 8), denn Lernerkorpora weisen gewisse Vorteile gegenüber traditionellen Datensammlungen in der Z/FSE-Forschung auf. Zu den traditionellen Datensammlungen in der Z/FSE-Forschung zähle ich an dieser Stelle vorrangig Fehlersammlungen (vgl. Fries 1988, Clyne 1968) oder Einzelfallstudien. Letztere dominieren in der deutschsprachigen Z/FSE-Forschung (Hufeisen/Riemer 2010: 748). Der größte Vorteil von Lernerkorpusdaten gegenüber Sprachdaten aus Einzelfallstudien ist die Datenmenge. Lernerkorpora bieten i.d.R. die notwendige quantitative Basis zur Überprüfung aktueller theoretischer Annahmen zum Spracherwerb und zur empirischen Validierung von Erkenntnissen aus bisherigen Spracherwerbsuntersuchungen auf Basis geringerer Sprachdaten (Granger 2008: 269). Somit sind die in einem Lernerkorpus enthaltenen Daten im Vergleich zu traditionellen Datensammlungen in der Z/FSE-Forschung repräsentativer, da sie i.d.R. Sprachdaten vieler verschiedener Lernenden enthalten (Granger et al. 2015: 1). Aus korpuslinguistischer Sicht betrachtet sind Lernerkorpora nach Goschler/Stefanowitsch (2014: 346) sehr klein und enthalten Sprachdaten „einiger weniger Lernender“, sodass individuelle Besonderheiten bei Analysen ihrer Ansicht nach zu stark gewichtet werden¹¹⁹. Verglichen mit traditionellen Datensammlungen in der Z/FSE-Forschung, in denen z. B. zwei (Kuhberg 1990), neun

¹¹⁸ Zum Verhältnis von generativer Grammatik und Korpuslinguistik im Allgemeinen siehe z. B. Lemnitzer/Zinsmeister (2015: 22ff.).

¹¹⁹ Goschler/Stefanowitsch (2014: 348ff.) merken zudem an, dass die Untersuchung von Lernaltersprache auf Basis von Lernerkorpora ein Problem darstellt, da Lernaltersprache durch Individualität, Instabilität und Variabilität geprägt ist. Dies scheint mir jedoch kein Problem von Lernerkorpora zu sein, sondern eine Besonderheit des Wesens von Lernaltersprache im Allgemeinen, unabhängig davon, auf welcher Datengrundlage Lernaltersprache empirisch untersucht wird. Das Problem ist auch kein Grund, vollständig auf Lernerkorpora in der Z/FSE-Forschung zu verzichten, das räumen auch die Autoren abschließend ein (Goschler/Stefanowitsch 2014: 356).

(Wegener 2008) oder elf (Baten/Willems 2012) Lernende die Datengrundlage bilden, beinhalten Lernerkorpora dennoch Daten von deutlich mehr Lernenden, z. B. ca. 200 Lernenden (Falko, Reznicek et al. 2012: 29) oder 1000 Lernenden (MERLIN, Merlin-Projekt 2014: 26).

Im Vergleich zu bloßen Fehlersammlungen eignen sich Lernerkorpora besser für die Spracherwerbsforschung,

da der Kontext einer verwendeten Äußerung mitbetrachtet werden kann, sowohl ‚abweichende‘ und ‚zielsprachliche‘ Äußerungen desselben Lernalters bei einer bestimmten Konstruktion untersucht werden können und quantitative Vergleiche zu [muttersprachlichen] Daten oder Lernerdaten anderer Lernergruppen (im Längsschnitt oder im Querschnitt) möglich sind (Lüdeling/Walter 2010: 318).

Lernerkorpora bieten somit den Vorteil, dass sie vollständige Produktionen von Sprachlernenden zur Analyse bereitstellen und nicht nur Sprach-Fragmente abbilden.

Ein weiterer Vorteil von Lernerkorpora liegt in der Anwendung korpuslinguistischer Werkzeuge und Analysemethoden. Die prototypische Lernerkorpusforschung greift auf (teil)automatische Analyse- und Annotationswerkzeuge zurück (Rankin 2015: 236) oder bedient sich im Falle von z. B. Fehlerannotation entsprechender unterstützender Annotationssoftware. Die Anwendung derartiger Werkzeuge und Software beschleunigt die Datenanalyse bzw. macht eine solche erst möglich (Granger et al. 2015: 1f.).

Neben den bisher genannten Vorteilen von Lernerkorpusdaten möchte ich an dieser Stelle noch einen weiteren Vorteil nennen, der die Dokumentation der Daten bzw. die Dokumentation der Datenerhebung, -aufbereitung und -auswertung betrifft. Rankin (2015: 236) kontrastiert die methodologischen Ansätze der Lernerkorpusforschung an der Schnittstelle zur Z/FSE-Forschung mit denjenigen der traditionellen Z/FSE-Forschung am Beispiel von Untersuchungen zur Morphosyntax. Er erwähnt, dass die traditionelle Z/FSE-Forschung selten die Datenerhebung, -aufbereitung und -analyse explizit in ihren Publikationen diskutiert. Ähnliches stellen auch Gass/Selinker (2008: 57) fest. Sie erwähnen, dass in Publikationen der Z/FSE-Forschung oft keine detaillierte Beschreibung der Probanden oder der Datenerhebung zu finden seien. Dies ist jedoch keine Kritik an der entsprechenden Forschung, so Gass/Selinker (2008: 57), sondern an den Darstellungen der Untersuchungsergebnisse in entsprechenden Publikationen. Diese Art der Darstellung erschwert die Nachvollziehbarkeit wissenschaftlicher Ergebnisse (Gass/Selinker 2008: 57). Zudem erschwert eine solche Art der Darstellung die Replizierbarkeit wissenschaftlicher Untersuchungen. Im Gegensatz dazu findet man in korpuslinguistischen Publikationen an der Schnittstelle zur Z/FSE-Forschung eine explizite Darstellung des entsprechenden Lernerkorpus, der darin enthaltenen Daten und Metadaten. Ebenfalls wird die Datenaufbereitung und -analyse explizit thematisiert und reflektiert (vgl. z. B. Arbeiten von Breckle 2014, Walter/Doolittle 2012). Wissenschaftliche Erkenntnisse sind somit besser nachvollziehbar und replizierbar; letzteres jedoch nur unter der Voraussetzung, dass das jeweilige Lernerkorpus für Forschungszwecke (frei) zugänglich ist.

In Publikationen der Lernerkorpusforschung stellt oft auch die Korpusmethodologie selbst den Untersuchungsgegenstand dar, so wie z. B. bei Lüdeling et al. (2005), die die Anforderungen an eine Lernerkorpusarchitektur unter Berücksichtigung der Charakteristiken lernersprachlicher Produktionen thematisieren oder bei Lüdeling (2008), die die Mehrdeutigkeiten bei der Fehlerkorrektur und -kategorisierung in lernersprachlichen Produktionen und die Auswirkungen auf die Fehlerannotation und -analyse diskutiert.

Aufgrund der genannten Vorteile wird Lernerkorpora mittlerweile eine „besondere Rolle“ in der empirischen Z/FSE-Forschung zugeschrieben (K. Schmidt 2010: 555). Ihr Potenzial für die Spracherwerbsforschung (auch im DaF-Bereich) ist bekannt (Granger 2002, 2009, 2012, K. Schmidt 2010) und hat in den vergangenen Jahren an Relevanz zugenommen. Sprechen Walter/Grommes (2008: 15) von Lernerkorpora noch als "ausgewöhnliche Datenbasis" in der Z/FSE-Forschung, erwähnen Lüdeling/Walter (2010), dass „immer häufiger Lernerkorpora als Datenbasis für die Erwerbsforschung eingesetzt werden“. K. Schmidt (2010) macht deutlich, dass „Spracherwerbsforscher im Bereich Deutsch als Fremdsprache sich zukünftig mit den Möglichkeiten und Grenzen von korpuslinguistischen Analysen [...] von Lernerkorpora [...] auseinandersetzen müssen" (558f.). Goschler/Stefanowitsch (2014: 342f.) stellen ebenfalls fest, dass „(Zweitspracherwerbs)korpora und korpuslinguistische Methoden neben anderen empirischen Zugängen zum Zweitspracherwerb schon lange zu den anerkannten empirischen Grundlagen der Zweitspracherwerbsforschung [gehören]". Im Kontext der Z/FSE-Forschung gewinnen digitale Lernerkorpora als Sprachressource somit immer mehr an Bedeutung.

Trotz der Überlappungen im Erkenntnisinteresse der traditionellen Z/FSE-Forschung und der Lernerkorpusforschung an deren Schnittstelle sowie der Vorteile der lernerkorpuslinguistischen Methodologie gegenüber traditionellen Datensammlungen der Z/FSE-Forschung nähern sich beide Forschungsfelder nur schrittweise an (Granger 2012: 8). Das Potenzial von Lernerkorpora für die Z/FSE-Forschung wird zwar vielerorts betont, Lernerkorpora gelten innerhalb der Z/FSE-Forschung aber immer noch als eine relativ neue Datengrundlage (Granger 2012). Ein Hauptgrund dafür besteht darin, dass die Lernerkorpusforschung an der Schnittstelle zur Z/FSE-Forschung eher deskriptiv statt interpretativ ausgerichtet und immer noch zu wenig theoriegeleitet ist (Myles 2015: 309, Rankin 2015: 239), auch wenn sich aktuell Tendenzen zur theoriegeleiteter Lernerkorpusforschung an der Schnittstelle zur Z/FSE-Forschung beobachten lassen (vgl. z. B. Lüdeling/Hirschmann/Shadrova 2017). Die (bisher noch) mangelnde Orientierung an spracherwerbstheoretischen Erkenntnissen lässt sich sicherlich primär darauf zurückführen, dass die Lernerkorpusforschung stark von der Korpuslinguistik und damit von der sprachwissenschaftlichen Forschung geprägt ist, die ebenfalls einen deskriptiven Charakter hat.

Ein weiterer Grund, warum sich die Lernerkorpusforschung und die traditionelle Z/FSE-Forschung nur langsam annähern, sind die typischerweise in einem Lernerkorpus verfügbaren Daten. Die meisten Lernerkorpora enthalten schriftliche Daten (vorwiegend argumentative Texte) von fortgeschrittenen Lernenden (Querschnitt) einer Zielsprache, wobei die Anzahl der repräsentierten Erstsprachen i.d.R. begrenzt ist (Gilquin 2015: 12f.). Die traditionelle Z/FSE-Forschung ist jedoch vorwiegend interessiert an mündlichen Daten (aus diversen

kommunikativen Kontexten), unterschiedlichen Kompetenzniveaus (Längsschnitt) und Lernenden, die unterschiedliche Erstsprachen repräsentieren (Myles 2015: 316). Dies sind auch die ‚Lücken‘, die nach Myles (2015: 328ff.) in der Lernerkorpusforschung geschlossen werden müssen, damit die Lernerkorpusforschung und die traditionelle Z/FSE-Forschung erfolgreich in einen Dialog miteinander treten und (mehr) voneinander profitieren können. Gleichzeitig macht Myles (2015: 330) auf die unterschiedlichen Ansätze zum Erkenntnisgewinn aufmerksam, die einen vollständigen Dialog zwischen beiden Forschungsfeldern und damit einhergehend die Nutzung von Lernerkorpora in der Z/FSE-Forschung (noch) erschweren. Nach Myles (2015: 330) sind die aktuell verfügbaren Lernerkorpora¹²⁰ gut geeignet für bottom-up-Ansätze der Datenanalyse, d. h. für die Ermittlung sprachlicher Muster und Tendenzen in einer Vielzahl von Lernerdaten. Dahingegen bevorzugt die traditionelle Z/FSE-Forschung einen top-down-Ansatz, also die Analyse von Lernerdaten zur Überprüfung bestimmter theoretischer Annahmen zum Spracherwerb. Nach Goschler/Stefanowitsch (2014: 356) eignen sich jedoch Lernerkorpora ebenfalls sehr gut für hypothesenüberprüfende Untersuchungen und sollten für eben solche Untersuchungen stärker eingesetzt werden.

Aktuelle Entwicklungen in der Lernerkorpusforschung lassen eine wachsende Kooperation beider Forschungsfelder feststellen: Einerseits wird das Lernerkorpus im Kontext der Z/FSE-Forschung als Datengrundlage thematisiert (Granger 2012, Walter/Grommes 2008) und andererseits ist z. B. im aktuellen Handbuch zur Lernerkorpusforschung (Granger/Gilquin/Meunier 2015) ein thematischer Komplex integriert, der sich explizit der Schnittstelle zur traditionellen Z/FSE-Forschung widmet. In unterschiedlichen Beiträgen wird dort das Potenzial von Lernerkorpora für die Z/FSE-Forschung thematisiert, so z. B. für die Erforschung von Entwicklungsverläufen im Spracherwerb (Meunier 2015). Ein Austausch zwischen beiden Forschungsfeldern wird ebenfalls gefördert durch die Learner Corpus Association, einem 2013 gegründetem Netzwerk zur Förderung der Lernerkorpusforschung und des interdisziplinären Austauschs an ihren Schnittstellen. Abschließend lässt sich festhalten, dass trotz aktueller bidirektionaler Bewegungen zwischen der Lernerkorpusforschung und der traditionellen Z/FSE-Forschung derzeit aufgrund unterschiedlicher Anforderungen an die Datenbasis sowie unterschiedlichen Ansätzen der Datenanalyse in beiden Forschungsfeldern (vgl. Myles 2015: 330, s. oben) noch weiterer ‚Annäherungsbedarf‘ besteht.

Zum Abschluss möchte ich auf ausgewählte Publikationen verweisen, die den Einsatz von deutschen Lernerkorpora¹²¹ an der Schnittstelle zur Z/FSE-Forschung illustrieren. Die folgende Tabelle (Tab. 15) gibt einen entsprechenden Überblick (alphabetisch sortiert).¹²²

¹²⁰ Myles (2015) erwähnt lediglich Lernerkorpora der Zielsprache Englisch.

¹²¹ Ich beschränke mich hier auf Lernerkorpora, die schriftliche Daten von Deutschlernenden enthalten. Für ein Lernerkorpus mit mündlichen Daten von Deutschlernenden siehe Gut (2012). Eine Übersicht zu verschiedenen Lernerkorpora (u.a. Lernerkorpora der Zielsprache Deutsch) bietet die Seite *Learner corpora around the world* (LCW).

¹²² Die vollständigen Literaturangaben sind im Literaturverzeichnis dieser Arbeit zu finden.

Publikation Kurzverweis	Untersuchungsgegenstand	Lernerkorpus
Breckle (2014)	Pronomen	Kobalt-Korpus
Breckle/Zinsmeister (2010)	Definite/indefinite Artikel	ALeSKO
Breckle/Zinsmeister (2012)	Wortstellung (Vorfeld, pragmatische Funktion des Vorfelds)	ALeSKO
Breckle/Zinsmeister (2013)	Wortstellung (Vorfeld, Chunks)	ALeSKO
Breindl (2013)	Präposition-Artikel-Verschmelzungen	Falko Essay ¹²³
Breindl (2016, 2018)	Konnektoren	Kobalt-Korpus
Hirschmann et al. (2013)	Modifikatoren	Falko Essay
Lüdeling/Hirschmann/Shadrova (2017)	Komplexe Verben (Präfix- und Partikelverben)	Falko Essay
Maden-Weinberger (2008)	Modalität (Modalverben/modale Adverbiale ¹²⁴)	CLEG
Maden-Weinberger (2015)	Konjunktiv II	CLEG
Sieradz/Bordag (2014)	Präposition-Artikel-Verschmelzungen	Falko Essay
Walter/Doolittle (2012)	Wortstellung (Vorfeld, Syntaktische Komplexität)	Falko GU ¹²⁵
Wisniewski (2018b)	Grammatische Fehler (in versch. Kompetenzniveaus)	MERLIN

Tabelle 15: Ausgewählte Publikationen (deutsche Lernerkorpora) an der Schnittstelle zur Zweit- und Fremdspracherwerbsforschung.

In drei der in Tabelle 15 aufgeführten Publikationen lässt sich ein expliziter Bezug zu Spracherwerbstheorien finden (Breckle/Zinsmeister 2010, Maden-Weinberger 2015, Lüdeling/Hirschmann/Shadrova 2017). Die restlichen Publikationen haben, wie auch die Mehrheit der Untersuchungen in deutschen Lernerkorpora an der Schnittstelle zur Z/FSE-Forschung, einen deskriptiven Schwerpunkt. Das Lernerkorpus MERLIN wird bisher (Stand März 2019) kaum in der deutschen Lernerkorpusforschung eingesetzt, obwohl dieses durch seine reliabel ermittelten Kompetenzniveaus (vgl. Wisniewski 2018b: 134) vor allem ein großes Potenzial für Forschungen zu Erwerbsverläufen birgt.

Die korpusgestützten Fallstudien zu ausgewählten Einflussvariablen im Gebrauch von Präpositionen und Präpositionalphrasen durch Deutschlernende präsentieren einen empirischen Zugang, der aktuelle Annahmen der Zweit- und Fremdspracherwerb auf Basis von Lernerkorpusdaten überprüft. Mit der vorliegenden Dissertation wird somit das Potenzial von Lernerkorpora als Ressource für die quantitative Z/FSE-Forschung veranschaulicht.

¹²³ Das Falko Korpus besteht u.a. aus einem Subkorpus mit argumentativen Texten (Falko Essay), vgl. dazu Lüdeling et al. (2008).

¹²⁴ Zu der Gruppe der modalen Adverbialien zählt Maden-Weinberger (2008: 149) u.a. Adjektivphrasen wie *voraussichtlich*, PP wie *mit Sicherheit* und verbale Konstruktionen wie *es ist anzunehmen*.

¹²⁵ GU steht für Georgetown University. Für weitere Informationen s. Walter/Doolittle (2012: 3f.).

3.6 Zusammenfassung

Die Lernerkorpusforschung beschäftigt sich mit der Analyse von Lernerdaten in einem Lernerkorpus – einer systematischen, digitalen und maschinenlesbaren Sammlung authentischer sprachlicher Produktionen von Sprachlernenden. In der Regel sind in einem Lernerkorpus Lernende einer Fremdsprache vertreten; ‚authentische‘ Produktionen meinen in diesem Zusammenhang primär eine authentische Sprachlernsituation, in deren Rahmen die in einem Lernerkorpus enthaltenen Daten erhoben werden.

Lernerkorpora kann man entlang unterschiedlicher Dimensionen in unterschiedliche Typen klassifizieren. Gemäß der Typologie, die in diesem Kapitel vorgestellt wurde, kann man feststellen, dass ein aktuell durchschnittliches Lernerkorpus schriftliche Daten (i.d.R. argumentative Texte) von Lernenden des Englischen als Zielsprache enthält, die nicht weiter aufbereitet, d. h. annotiert sind. Weiterhin stammen die Daten solch eines aktuell durchschnittlichen Lernerkorpus von Lernenden mit einer gemeinsamen Erstsprache und wurden im Rahmen eines Forschungsprojekts in einer Querschnittserhebung gesammelt, um für wissenschaftliche und ggf. pädagogische Zwecke eingesetzt zu werden. Die internationale Lernerkorpuslandschaft entwickelt sich jedoch stetig weiter. So gibt es u.a. mittlerweile auch annotierte Lernerkorpora mit anderen Zielsprachen als Englisch, in denen Sprachlernende mit verschiedenen Erstsprachen und Kompetenzniveaus repräsentiert sind. Zu solchen Lernerkorpora gehört auch das Lernerkorpus MERLIN, das als Datengrundlage der Fallstudien in der vorliegenden Dissertation dient (Näheres zu MERLIN s. Kap. 6). Eine umfassende Liste von *Learner corpora around the world* (LCW) bietet das Centre for English Corpus Linguistics der Universität Louvain¹²⁶.

Die Lernerkorpusforschung liegt an den Schnittstellen zu unterschiedlichen Forschungsfeldern. Die zentrale Schnittstelle ist die Schnittstelle zur Zweit-/Fremdspracherwerbsforschung (Z/FSE-Forschung). Die ersten Lernerkorpora waren durch den Transfer und die Anwendung korpuslinguistischen Wissens und korpuslinguistischer Methoden motiviert; Ziel war (und ist es immer noch), ein besseres Verständnis des Wesens der Lernaltersprache und ihrer Entwicklung zu ermöglichen sowie die Faktoren, die Wesen und Entwicklung der Lernaltersprache beeinflussen, zu ermitteln – eine Zielsetzung, die ebenfalls die traditionelle Z/FSE-Forschung verfolgte (und immer noch verfolgt). Lernerkorpusdaten sind vor allem geeignet im Kontext kognitiver Spracherwerbtheorien und im Rahmen von Untersuchungen zu Merkmalen des linguistischen Systems der Lernaltersprache sowie zu Einflussvariablen im Sprachgebrauch. Die Vorteile von Lernerkorpora gegenüber traditionellen Datensammlungen der Z/FSE-Forschung, nämlich Einzelfallstudien und Fehlersammlungen, liegen vor allem in der Quantität (und damit Repräsentativität) der Daten sowie den Möglichkeiten einer (teil)automatischen Extraktion und Analyse relevanter sprachlicher Strukturen. Lernerkorpora haben in den vergangenen Jahrzehnten an Relevanz für die Z/FSE-Forschung gewonnen, dennoch nähern sich die beiden Forschungsfelder nur schrittweise an. Die Z/FSE-Forschung muss das Potenzial von Lernerkorpora nutzen, die

¹²⁶ <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html> (Mai 2019).

Lernerkorpusforschung dahingegen stärker spracherwerbtheoretische Erkenntnisse in ihren Datenanalysen berücksichtigen.

Die vorliegende Dissertation zeigt am Beispiel von Präpositionen und PP, wie zentrale Annahmen kognitiver Spracherwerbtheorien anhand von Lernerkorpusdaten quantitativ überprüft und mögliche Einflussvariablen im Sprachgebrauch von DaF-Lernenden ermittelt werden können. Auf diese Weise bettet sich die vorliegende Arbeit in die Lernerkorpusforschung an der Schnittstelle zur Zweit- und Fremdspracherwerbsforschung ein und illustriert das Potenzial von Lernerkorpora als Ressource für die quantitative DaF-Spracherwerbsforschung.

4. Skizzierung des Forschungsstands

4.1 Präpositionen und PP in der deutschen Zweit- und Fremdspracherwerbsforschung

In diesem Kapitel stelle ich den Forschungsstand zum Erwerb und Gebrauch von Präpositionen und Präpositionalphrasen (PP) durch Lernende des Deutschen aus dem Bereich der Zweit- bzw. Fremdspracherwerbsforschung dar. Der Fokus der Darstellungen bisheriger Erkenntnisse liegt dabei primär auf dem Gebrauch der Präposition und der entsprechenden PP. Dabei lasse ich Erkenntnisse zur Kasusreaktion aus, da die Kasusreaktion in den empirischen Untersuchungen der vorliegenden Dissertation nicht untersucht wird¹²⁷. Weiterhin liegt der Schwerpunkt dieses Kapitels auf Erkenntnissen zur Produktion von Präpositionen und PP. Forschungsarbeiten, die sich auf die Rezeption von Präpositionen und PP beziehen, werden nicht thematisiert (s. hierzu z. B. Gürsoy/Wilhelm 2014). Die in diesem Kapitel erwähnten Forschungserkenntnisse stammen sowohl aus Untersuchungen zum ungesteuerten Spracherwerb von Kindern und Erwachsenen als auch aus Untersuchungen zum gesteuerten Spracherwerb Erwachsener, wobei zu erwachsenen Sprachlernenden nach Rösch (2011: 11) grundsätzlich alle Personen jenseits der Pubertät gezählt werden. Ich folge Rösch (2011: 16) und verstehe den Erwerb des Deutschen als Zweitsprache (DaZ-Erwerb) als den weitestgehend ungesteuerten (schulischen) Erwerb der deutschen Sprache im Zielland, bei dem die Zielsprache Deutsch als *Instruktionssprache* fungiert. Unter dem Erwerb des Deutschen als Fremdsprache (DaF-Erwerb) verstehe ich dagegen in Anlehnung an Rösch (2011: 16) den gesteuerten Erwerb der deutschen Sprache, bei dem die Zielsprache Deutsch den *Lerngegenstand* darstellt.

Ziel dieses Kapitels ist die Herausarbeitung der zentralen Erkenntnisse und Forschungsdesiderate zu (potentiellen)¹²⁸ Einflussvariablen im Gebrauch von Präpositionen und Präpositionalphrasen (PP) durch Lernende des Deutschen (DaZ und DaF), die für die

¹²⁷ Für Untersuchungen, die sich vorrangig mit der Kasusreaktion nach Präpositionen befassen, siehe z. B. Baten/Willems (2012), Wilmots/Moonen (1997), Turgay (2010b), Turgay (2011), Ott (2006), Freitag/Vandermeeren (2005) und Kwakernaak (2005).

¹²⁸ Das Wort *potentiell* ist hier in Klammern, da im Forschungsüberblick sowohl Erkenntnisse zu gut erforschten Einflussvariablen im Präpositionsgebrauch präsentiert werden als auch zu solchen Einflussvariablen, bei denen bisher nicht ausreichend empirische Erkenntnisse vorliegen.

empirischen Untersuchungen dieser Dissertation von Relevanz sind. Obwohl in der vorliegenden Dissertation ausschließlich Produktionen von DaF-Lernenden untersucht werden (vgl. MERLIN-Projekt 2014: 3), werden in diesem Kapitel sowohl Erkenntnisse aus dem DaZ- als auch dem DaF-Bereich berücksichtigt. Damit werden auch Erkenntnisse zu Deutschlernenden unterschiedlichen Alters abgedeckt. Rösch folgend gibt es jedoch „vielfältige Parallelen zwischen dem Zweitspracherwerb von Kindern, Jugendlichen und Erwachsenen“ (Rösch 2011: 11). Durch die Berücksichtigung von Erkenntnissen aus dem DaZ- als auch dem DaF-Bereich gebe ich einen umfassenden Überblick über den bisherigen Forschungsstand.

Das vorliegende Kapitel präsentiert jeweils zentrale Forschungserkenntnisse hinsichtlich der (potentiellen) Einflussvariablen im Erwerb und Gebrauch von Präpositionen und PP und ist entsprechend dieser Einflussvariablen gegliedert. Die Einflussvariablen sind: Die Erstsprache der Lernenden (Kap. 4.1.2), die Sprachkompetenz der Lernenden (Kap. 4.1.3), die syntaktische Funktion der PP (Kap. 4.1.4) sowie der semantische Gehalt der Präposition (Kap. 4.1.5). Kapitel 4.1.6. bietet eine Gesamt-Zusammenfassung der Forschungserkenntnisse.

Im Folgenden werde ich zunächst einen Überblick zu den Untersuchungen geben, die in diesem Forschungsüberblick berücksichtigt werden. Daran anschließend stelle ich die bisherigen zentralen Erkenntnisse der Zweit- und Fremdspracherwerbsforschung im Hinblick auf die oben genannten Einflussvariablen im Gebrauch von Präpositionen und PP nacheinander dar. Je Einflussvariable werden die Erkenntnisse in folgender Reihenfolge präsentiert: Zunächst werden Erkenntnisse aus der DaZ-Forschung dargelegt, wobei hier bei den Lernenden zwischen Kindern und Erwachsenen getrennt wird; daran anschließend werden Erkenntnisse aus der DaF-Forschung präsentiert. Am Ende des Kapitels fasse ich die zentralen Erkenntnisse zusammen.

4.1.1 Überblick

Die einzelnen Untersuchungen zum Gebrauch von Präpositionen und PP identifizierten unterschiedliche Herausforderungen und Lernschwierigkeiten, die sich für Deutschlernende zeigen. Dabei steht außer Frage, dass Präpositionen und PP eine Lernschwierigkeit für Deutschlernende darstellen (vgl. z. B. Steinmüller 1992: 28, Apeltauer 1987: 222). Ein Großteil der hier herangezogenen Untersuchungen widmet sich der Analyse des korrekten und vor allem inkorrekten Gebrauchs von Präpositionen und PP durch Deutschlernende. Die folgende Tabelle (Tab. 16) bietet einen Überblick über die empirischen Untersuchungen, die im vorliegenden Kapitel herangezogen werden.

Kurzverweis	Untersuchungsgegenstand	DaZ/DaF
Bryant (2012) Lütke (2008)	Präpositionen und PP = Teilbereich von Untersuchungen zu Raumreferenzen	DaZ (Kinder)
Becker (2012) Carroll/Becker (1993)		DaZ (Erwachsene)
Ahrenholz (1987) Harnisch (1993) Hepsöyler/Liebe-Harkort (1991) Kaltenbach/Klages (2006) Kuhberg (1990) Pfaff (1984) Pienemann (1981)	Präpositionen und PP = Teilbereich von Untersuchungen zum Erwerb und Gebrauch sprachlicher Strukturen in der Zielsprache Deutsch	DaZ (Kinder)
Balci (2006) Clyne (1968) Diehl et al. (2000) HPD ¹²⁹ (1977) Keim (1984) Meisel (1977) Pepouna et al. (2018) Roe (2007)		DaZ (Erwachsene)
Becker (2006) Grießhaber (2007b) Gutzmann/Turgay (2011) Kalkavan-Aydin (2017) Lütke (2010), (2011) Turgay (2010a, b, 2011)	Präpositionen und PP = Fokus der Untersuchung	DaZ (Kinder)
Bouillon (1977) Gibson/Hufeisen/Libben (2001) Hawthorne (2005) ¹³⁰ Hufeisen/Gibson (2002) Viorel (1973)		DaF (Erwachsene)

Tabelle 16: Untersuchungen zu Präpositionen und PP im DaZ/DaF-Erwerb: Überblick.

Im Folgenden stelle ich die zentralen Erkenntnisse zu (potentiellen) Einflussvariablen im Gebrauch deutscher Präpositionen und PP vor, die sich auf Basis bisheriger empirischer Untersuchungen ableiten lassen.

¹²⁹ Die Abkürzung HPD steht hier und im Folgenden für die Publikation der Forschungsergebnisse des Heidelberger Forschungsprojekts *Pidgin-Deutsch spanischer und italienischer Arbeiter in der BRD*.

¹³⁰ Sowohl Gibson/Hufeisen/Libben (2001), Hufeisen/Gibson (2002) als auch Hawthorne (2005) arbeiten mit Aufgaben, die ein Einsetzen von Präpositionen in vordefinierten Kontexten erfordern. Bei diesen Untersuchungen liegt somit eine kontrollierte Datenerhebung vor.

4.1.2 Erstsprache der Lernenden

Eine insgesamt gut untersuchte Einflussvariable im Gebrauch von Präpositionen und PP durch Deutschlernende stellt die Erstsprache (L1) dar. Ihr Einfluss auf den Gebrauch von Präpositionen und PP im DaZ-Erwerb (bei Kindern) wird von Pfaff (1984), Kuhberg (1990), Griebhaber (2007b), Turgay (2010a), Gutzmann/Turgay (2011), Bryant (2012) und Kalkavan-Aydin (2017) explizit thematisiert und untersucht. Türkisch als L1 steht dabei im Vordergrund – in allen eben aufgeführten Untersuchungen werden Daten von Lernenden mit L1 Türkisch analysiert¹³¹. In der Mehrheit der Fälle werden auch weitere L1 untersucht bzw. der L1 Türkisch gegenübergestellt, so z. B. das Italienische (Turgay 2010a), das Russische (Bryant 2012), das Griechische (Pfaff 1984) oder das Polnische (Kuhberg 1990). Griebhaber (2007b) untersucht Daten von Kindern mit den Erstsprachen Türkisch, Serbokroatisch, Griechisch, Indonesisch, Albanisch und Arabisch. Dabei stehen die Kinder mit L1 Türkisch im Fokus und werden mit den Kindern anderer L1 kontrastiert.

Turgay (2010a) und Bryant (2012) stellen einen möglichen Einfluss der L1 auf das Erwerbtempo von Präpositionen und PP fest. So zeigt Turgay (2010a: 278) in der Analyse lernersprachlicher Daten, dass Kinder mit L1 Türkisch ein langsames Erwerbtempo aufweisen als Kinder mit L1 Italienisch, da türkischsprachige Kinder später Präpositionen produzieren als italienischsprachige Kinder. Die korrekte Wahl einer Präposition wird ihren Daten nach von Deutschlernenden mit L1 Türkisch allerdings schneller beherrscht als von Deutschlernenden mit L1 Italienisch (Turgay 2010a: 278). Turgay (2010a: 279) kommt zum Schluss, dass Lernende mit L1 Türkisch u.a. in Bezug auf die Präpositionswahl ein schnelleres Erwerbtempo zeigen als Lernende mit L1 Italienisch; Lernende mit L1 Italienisch dahingegen weisen u.a. ein schnelleres Erwerbtempo in Bezug auf die Realisierung von Präpositionen und der entsprechenden PP auf. Diese Beobachtungen führt Turgay (2010a) auf die Strukturen der jeweiligen L1 zurück, die den Erwerb von deutschen Präpositionen und PP beeinflussen – im Falle des Italienischen wird das Erwerbtempo positiv beeinflusst, da italienische PP viele Gemeinsamkeiten mit deutschen PP aufweisen (Turgay 2010a: 279). Turgay (2010a: 279) räumt jedoch ein, dass ihre Ergebnisse lediglich Tendenzen aufzeigen und sich in ihren Daten somit kein eindeutiger Einfluss der L1 zeigt. Weiterhin beobachtet Turgay (2010a: 273), dass Lernende mit L1 Türkisch keinen negativen Transfer bei der Position von Präpositionen zeigen, obwohl die türkische Sprache über Postpositionen verfügt, bei denen ein positionsbedingter Transfer in die Zielsprache Deutsch erwartet werden könnte. Insgesamt gestaltet sich der Gebrauch von Präpositionen und PP für beide Sprachgruppen (L1 Italienisch und L1 Türkisch) als schwierig, so Turgay (2010a: 276). Einen Einfluss der L1 auf den Gebrauch einzelner Präpositionen in der Zielsprache Deutsch stellen auch Kaltenbacher/Klages (2006: 88f.) im Ansatz fest. Sie führen dies zurück auf unterschiedliche Raumkonzepte und Versprachlichungen in L1 und Zielsprache.

¹³¹ Die Berücksichtigung von Kindern mit Türkisch als L1 ist hauptsächlich dadurch bedingt, dass diese Sprachgruppe zur größten Migrantengruppe in Deutschland zählt (vgl. Darstellungen in Turgay 2010a: 2f.). Daraus ergibt sich somit ein relativ leichter Zugang zu dieser Sprachgruppe sowie das Interesse an der Erforschung des Spracherwerbs und der Sprachentwicklung von Kindern dieser Sprachgruppe und den damit verbundenen sprachdidaktischen Implikationen.

Einen Einfluss der sprachlichen Strukturen der L1 auf die Phasen im Erwerb von deutschen Präpositionen und PP als Lokalisierungsausdrücke kann Bryant (2012) in Ihrer Untersuchung von Deutschlernenden mit L1 Türkisch und L1 Russisch beobachten: So scheinen die Lernenden mit L1 Türkisch ausgehend von den strukturellen Charakteristiken ihrer L1 „auf postpositionale betonte lokale Informationsträger“ (Bryant 2012: 228) zu achten. Dies führt dazu, dass die Lernenden mit L1 Türkisch in einem bestimmten Erwerbsstadium zwar zielsprachlich nicht korrekte und unspezifische Präpositionen (z. B. *in/bei*) verwenden, diese jedoch mit einem postpositionalen Element (Präpositionaladverb) verknüpfen, das lokal spezifisch ist (z. B. *Der Brief is bein Umschlag drin*, Bryant 2012: 227). Die unspezifischen Präpositionen, die Lernende mit L1 Türkisch gebrauchen und übergeneralisieren, interpretiert Bryant (2012: 229) als Äquivalent der aus der L1 Türkisch bekannten, ebenfalls unspezifischen Kasusmarkierung zur Signalisierung einer Lokalität (Kasus Lokativ). Mithilfe der Konstruktion aus unspezifischer Präposition und postpositionalem Präpositionaladverb nähern sich die türkischsprachigen Lernenden den deutschen Strukturen und damit dem zielsprachlich korrekten Gebrauch von Präpositionen und PP an. Ein derartiges Muster im Gebrauch von Präpositionen und PP kann Bryant (2012: 228) bei Kindern mit L1 Russisch dahingegen nicht beobachten. Diese benötigen im Vergleich zu Lernenden mit L1 Türkisch jedoch eine längere Zeit, um Präpositionen und PP zielsprachlich korrekt zu gebrauchen (Bryant 2012: 232). Dies ist insofern überraschend, als das Russische dem Deutschen strukturell und konzeptuell nähersteht als das Türkische und somit eine schnellere Beherrschung der Präpositionswahl bei Lernenden mit L1 Russisch zu erwarten wäre. Diese Ergebnisse stehen jedoch im Einklang mit den entsprechenden Tendenzen, die auch Turgay (2010a, s.o.) beobachtet.

Dass Deutschlernende mit L1 Türkisch häufig unspezifische lokale Präpositionen wie *in* und *bei* verwenden und diese im Vergleich zu deutschsprachigen Lernenden häufiger gebrauchen und auch häufiger übergeneralisieren, stellt auch Kalkavan-Aydin (2017) fest. Wie Bryant (2012: 229) sieht auch Kalkavan-Aydin (2017: 384) die häufige Verwendung des Lokalsuffixes in der türkischen Sprache als möglichen Einflussfaktor auf die Übergeneralisierungen im Gebrauch der unspezifischen Präpositionen *in* und *bei* in der Zielsprache Deutsch. Wie aus Bryants (2012), Turgays (2010a) und Kalkavan-Aydins (2017) Ausführungen deutlich wird, übt die L1 der Lernenden einen Einfluss auf die Lernstrategien sowie das Erwerbtempo im Bereich Präpositionsgebrauch aus. Weiterhin scheinen auch Übergeneralisierungen bei Präpositionen durch die L1 beeinflusst zu sein. In anderen Untersuchungen finden sich weitere Hinweise auf den Einfluss der L1 auf einzelne Fehlertypen im Präpositionsgebrauch.

Grießhaber (2007b) analysiert schriftliche Erzählungen, die mit Bildvorlagen eliziert wurden und fokussiert bei seiner Auswertung die Realisierung von PP mit dem Bezugselement (*Straßen-)Seite*¹³². Bei dieser Analyse stellt Grießhaber (2007b) fest, dass Lernende mit L1 Türkisch mehr Fehler bei der Wahl der korrekten Präposition zeigen als Lernende mit anderen L1; dies zeigt sich ebenfalls in Verbindung mit anderen Bezugswörtern wie z. B.

¹³² Grießhaber (2007b) wählte das Bezugsobjekt (*Straßen-)Seite* aus, da dieses Bezugsobjekt in der Bildvorlage eine zentrale Rolle einnimmt (379).

Jugendherberge, Berg und *Straße* (Grießhaber 2007b: 383). Dieses Ergebnis kann z. T. auf abweichende, d. h. zielsprachlich nicht korrekte Konzeptualisierung des dargestellten Bezugsobjekts (z. B. bei *in die Straße* ist *Straße* als räumliches Objekt konzeptualisiert) aber auch auf abweichende Interpretation und damit verbundene abweichende Versprachlichung der Bildvorlage zurückgeführt werden (Grießhaber 2007b: 384). Weiterhin stellt Grießhaber (2007b: 387f.) in einer Analyse des Gebrauchs von Präfixverben (z. B. *ankommen*) die Tendenz fest, dass Lernende mit L1 Türkisch dazu neigen, ihnen aus ihrer L1 bekannte syntaktische Strukturen mit sprachlichen Mitteln des Deutschen nachzubilden und es aufgrund dessen z. B. zur Auslassung von Präpositionen kommen kann. So beobachtet Grießhaber (2007b: 387) „die Tendenz türkischer Schüler[Innen] zum ersatzweisen Gebrauch präfigierter Verben ohne Präposition anstelle präfigierter Verben mit Präposition“.

Eine vergleichsweise häufigere Auslassung von Präpositionen bei Lernenden mit L1 Türkisch (im Vergleich zu Lernenden mit L1 Griechisch) stellt auch Pfaff (1984) fest. Allerdings führt sie diesen nicht auf einen direkten Einfluss der L1 zurück. Dahingegen klassifiziert sie die Übergeneralisierung der Präposition *in*, die in der Gruppe der Lernenden mit L1 Griechisch häufiger vorkommt als bei Lernenden mit L1 Türkisch, als einen Transfer aus dem Griechischen, da die der deutschen Präposition *in* entsprechende griechische Präposition *se* ein breites Bedeutungsspektrum aufweist, das in die Zielsprache Deutsch transferiert wird (Pfaff 1984: 296). Nach Pfaff (1984) führt nämlich just die typologische Nähe von L1 und Zielsprache (L2) zu mehr – vor allem mehr wahlbedingten – Fehlern im Präpositionsgebrauch. Pfaff (1984) erklärt dies folgendermaßen: „[o]nly where L1 and L2 are structurally parallel can lexical transfer take place“ (Pfaff 1984: 296). Das bedeutet, dass wahlbedingte Fehler im Präpositionsgebrauch vor allem dann auftreten, wenn Präpositionen und PP aus der L1 bekannt sind, jedoch z. B. andere semantische Reichweiten haben als in der Zielsprache Deutsch, was wiederum einen negativen Transfer bedingen kann (s. auch Uhlisch 1995: 229). Eine typologisch vom Deutschen weiter entfernte L1 wie das Türkische dahingegen führt zu einem beschränkten Transfer aus der L1 und damit zu einem vergleichsweise geringen Anteil an wahlbedingten Fehlern im Präpositionsgebrauch. Damit sind diese Ergebnisse ähnlich zu denen von Turgay (2010a) und Bryant (2012) (s.o.). Die Ergebnisse von Pfaff (1984) stehen jedoch teilweise in Widerspruch zu Grießhabers (2007b) Ergebnissen, der einen höheren Anteil an wahlbedingten Fehlern im Präpositionsgebrauch bei Lernenden mit L1 Türkisch im Vergleich zu Lernenden mit anderen L1 (u.a. Griechisch) feststellt. Jedoch differenziert Grießhaber (2007b) in seinen Analysen lediglich zwischen Lernenden mit L1 Türkisch und Lernenden mit anderen L1, eine separate Analyse der jeweiligen anderen L1 erfolgt bei ihm nicht.

Lütke (2011) kommt in einem Forschungsüberblick zum Erwerb lokaler Präpositionen bzw. PP zum Schluss, dass die typologische Distanz zwischen L1 und L2 einen Einfluss auf den Erwerb und Gebrauch von Präpositionen und PP hat. Nach ihrer Einschätzung ist der Erwerb von Präpositionen und PP für Lernende mit einer vom Deutschen typologisch weit entfernten L1 (z. B. Türkisch) „deutlich schwieriger“ als für Lernende mit einer dem Deutschen typologisch nahen L1 und dies kann zu Interferenzen führen (Lütke 2011: 111). Was genau unter „deutlich schwieriger“ zu verstehen ist, führt Lütke (2011) jedoch nicht explizit aus. In Bezug auf den

ungesteuerten Erwerb von Raumreferenzen (u. a. durch PP) durch Erwachsene postulieren Carroll/Becker (1993: 144) (s. auch Becker 2012) einen Einfluss der L1 auf das Erwerbstempo von raumreferentiellen Ausdrücken, während sie keinen Einfluss auf die Abfolge der einzelnen von ihnen herausgearbeiteten Erwerbsphasen annehmen. Carroll/Becker (1993: 144) kommen hier u.a. zum Schluss, dass z. B. die L1 Italienisch im Hinblick auf den Erwerb von Raumreferenzen in der Zielsprache Deutsch als „lernförderlich“ einzustufen ist. Steinmüller (1992) stellt fest, dass Lernende mit L1 Türkisch vor allem Schwierigkeiten bei der Wahl einer zielsprachlich korrekten Präposition haben. Jedoch beobachtet er auch bei deutschen MuttersprachlerInnen Abweichungen in diesem Bereich und folgert daraus, dass es sich um keine Fehler handelt, die für Lernende mit L1 Türkisch spezifisch sind (Steinmüller 1992: 28). Damit schließt er einen Einfluss der L1 auf den Gebrauch von Präpositionen und PP zwar nicht aus, jedoch scheint Steinmüller (1992) in der inkorrekten Wahl einer Präposition keinen eindeutigen Einfluss der L1 zu erkennen. Auch Fries (1988: 331f.) vertritt die Meinung, dass Schwierigkeiten im Präpositionsgebrauch (u. a. Auslassungen oder die inkorrekte Wahl von Präpositionen) unabhängig von der L1 der Lernenden auftauchen, einen möglichen Einfluss dieser schließt jedoch auch er nicht vollständig aus. Bei Clyne (1968), der den Sprachgebrauch ausländischer ArbeiterInnen mit unterschiedlichen L1 untersucht, findet sich kein Hinweis auf einen möglichen Einfluss der L1 auf den Gebrauch von Präpositionen. Er erwähnt, dass „die Tendenzen im Deutschen der Gastarbeiter[Innen] ziemlich einheitlich“ sind (Clyne 1968: 137). Da Clyne (1968: 138) u.a. die Auslassung von Präpositionen als ein Merkmal „der ‚Gastarbeiter-Sprache‘“ beschreibt, wird hier implizit deutlich, dass er einen einheitlichen Sprachgebrauch, und damit einheitlichen Präpositionsgebrauch von Lernenden mit unterschiedlichen L1 annimmt.

Interferenzen aus der L1 kann dahingegen Keim (1984) in ihren Daten von erwachsenen Deutschlernenden mit L1 Türkisch feststellen. Keim (1984: 141) interpretiert dabei die Auslassung von Präpositionen in Objekt-PP (Beispiel: „*Kinderler ich immer gucken*“¹³³) als direkte Interferenz aus dem Türkischen, in dem Objekt-PP nicht existieren und ein Verb lediglich Kasusobjekte regiert: So gibt es in derartigen Fällen im Türkischen „kein Äquivalent der deutschen Präpositionalphrase“ (Keim 1984: 141). Keim (1984: 141) nimmt hier folglich an, dass beim Transfer der sprachlichen Strukturen aus der L1 es zur Auslassung der Präposition bzw. zu einer Realisierung eines Kasuskomplements in der Form einer Nominalphrase kommt. Jedoch schätzt Keim (1984: 141) den Ausfall von Präpositionen in adverbialen PP nicht als Interferenzerscheinung ein, da „besonders Verben, die im Deutschen eine Situativ- oder eine Richtungsergänzung [=PP-Ergänzung] verlangen, auch im Türkischen ein der deutschen Präpositionalphrase entsprechendes Äquivalent [erfordern], das meist als Suffix realisiert wird“. Folglich stuft Keim (1984) die Auslassung der Präposition in *Kinderler ich immer gucken* als direkten Einfluss der L1 Türkisch ein, die Auslassung in *mir wohnen Tunceli*¹³⁴ jedoch nicht¹³⁵. In manchen Fällen beobachtet Keim (1984: 142) ebenfalls, dass die

¹³³ Als zielsprachige Rekonstruktion nimmt Keim (1984: 141) hier *nach den Kindern sehe ich immer* an.

¹³⁴ Als zielsprachige Rekonstruktion nimmt Keim (1984: 141) hier *wir wohnen in Tunceli* an.

¹³⁵ Keims (1984) Argumentation ist meiner Ansicht nach an dieser Stelle nicht vollends nachvollziehbar. Sowohl bei Objekt-PP als auch bei adverbialen PP hat man es im Deutschen mit Strukturen zu tun, die

Lernenden innerhalb von deutschen Sätzen Kasussuffixe aus ihrer L1 transferieren, z. B. in *Küchede Essen kochen*, wobei das Suffix *-de* im Türkischen zur Markierung des Lokativs gebraucht wird. In solchen Fällen wird auf die aus der L1 bekannte morphologische Markierung der Raumrelation zurückgegriffen. Ähnliches gilt für Fälle, in denen die Anordnung der Elemente innerhalb einer deutschen PP aus dem Türkischen übernommen wird und es so zur Verwendung von Präpositionen in Poststellung kommt (Beispiel: *deutsche Frau mit heiraten*), wie es Keim (1984: 144) ebenfalls in ihren Daten beobachtet. In diesem Fall kann von Interferenzerscheinungen aus dem Türkischen im Bereich des Präpositionsgebrauchs gesprochen werden (vgl. Keim 1984: 142). Auch Meyer-Ingwersen (1975: 75f.) dokumentiert derartige Einzelfälle von Präpositionen in Poststellung (*ich Auto mit fahren*) bei Lernenden mit L1 Türkisch.

Balcı (2006) beschreibt am Beispiel von DaF-Lernenden mit L1 Türkisch Lernprobleme, die auf Unterschiede zwischen L1 und L2 im Bereich der Verbvalenz zurückzuführen sind und somit zu Fehlern führen, denen „selbst fortgeschrittene Lerner[Innen] nicht ausweichen können“ (Balcı 2006: 239). So führt Balcı (2006: 240) Fehler im Gebrauch von Objekt-PP auf den Einfluss der Valenzstrukturen aus der L1 zurück. Andererseits merkt Balcı (2006: 240) an, dass Lernprobleme bei Objekt-PP „eine Eigentümlichkeit beim Erwerb des Deutschen“ darstellen. Damit impliziert er folglich, dass Lern- und Gebrauchsschwierigkeiten in diesem Bereich unabhängig von der L1 zu erwarten sind.

Bouillon (1977: 48ff.) schließt auf Basis der von ihm analysierten Sprachdaten französischsprachiger DaF-Lernender, dass die strukturellen Unterschiede zwischen L1 (bzw. weiteren, bereits erlernten Sprachen) und der L2 Deutsch Interferenzerscheinungen im Präpositionsgebrauch bedingen. So führt Bouillon (1977: 48) die Wahl einer inkorrekten deutschen Präposition auf eine Übertragung bzw. Übergeneralisierung französischer Präpositionen zurück¹³⁶. Zu ähnlichen Annahmen kommt auch Viorel (1973) in ihrer Analyse einer entsprechenden Fehlersammlung von Deutschlernenden mit L1 Rumänisch: Fehler im Präpositionsgebrauch führt Viorel (1973: 354f.) vorrangig auf fehlende Übereinstimmungen im Präpositionssystem der L1 und der L2 und damit auf Interferenzen zurück. Die Verwendung inkorrekt hinzugefügter Präpositionen führen sowohl Viorel (1973) als auch Bouillon (1977) ebenfalls auf einen Einfluss der L1 zurück: So schreibt Bouillon (1977: 50), dass „Fehler dieser Art [...] [auf]treten, wenn im Deutschen keine Präposition vorhanden ist, wo in der Muttersprache aber in gleicher Position wohl eine Präposition vorhanden ist“. Einen ähnlichen Schluss zieht auch Viorel (1973: 355): Sie beobachtet in ihrer Fehlersammlung, dass inkorrekt hinzugefügte Präpositionen (z. B. *Ich habe auf die Großmutter gesehen*) in solchen Kontexten auftreten, in denen an entsprechender Stelle in der L1 Rumänisch eine Präposition gefordert

im Türkischen abweichend, und zwar i.d.R. mithilfe von Kasussuffixen realisiert werden, da PP nicht als Struktur zur Verfügung stehen. So werden PP im Deutschen in beiden Fällen i.d.R. durch NPs im Türkischen wiedergegeben.

¹³⁶ Beim Beispiel *Am zwei Uhr fährt er nach Köln* nimmt Bouillon (1977: 48) an, dass der entsprechende Lernende „wahrscheinlich *a deux heure* meint und *à* phonologisch mit *an* in Beziehung bringt“.

wäre – beide AutorInnen sehen folglich starke Einflüsse der L1 auf den Präpositionsgebrauch in der L2 Deutsch¹³⁷.

Dahingegen können Gibson/Hufeisen/Libben (2001) sowie Hufeisen/Gibson (2002) bei den von ihnen untersuchten DaF-Lernenden nur bedingt einen Einfluss bereits erlernter Sprachen auf den Gebrauch von deutschen Präpositionen und PP feststellen. Gibson/Hufeisen/Libben (2001) untersuchen die Kenntnis verbregierter Präpositionen (z. B. *warten auf*) bei 64 Lernenden, die Deutsch als erste oder zweite Fremdsprache gelernt haben¹³⁸. Die Ergebnisse ihrer Untersuchungen zeigen, dass die typologische Nähe der L1 zur L2 Deutsch keinen Einfluss auf die Korrektheit im Gebrauch verbregierter Präpositionen hat (Gibson/Hufeisen/Libben 2001: 143) – Lernende mit L1 Russisch schnitten ähnlich gut ab wie Lernende mit L1 Chinesisch (Gibson/Hufeisen/Libben 2001: 146). Des Weiteren deuten die Ergebnisse darauf hin, dass Lernende, die Englisch (als Fremdsprache) vor dem Deutschen erworben haben, eine leicht schlechtere Leistung bei Verben aufweisen, die im Englischen eine ähnliche Präpositionsreaktion wie im Deutschen zeigen (*gehören zu* – *belong to*) im Vergleich zu Verben, die keine ähnliche Präpositionsreaktion zeigen (*hoffen auf* – *hope for*). Die AutorInnen erklären dieses Ergebnis damit, dass just die Kenntnis englischer Verben bei den Lernenden möglicherweise zu einer negativen Interferenz geführt hat und damit zum fehlerhaften Gebrauch verbregierter Präpositionen im Deutschen: Die Lernenden sind sich der grundsätzlichen sprachlichen Unterschiede bewusst und vermeiden daher absichtlich den Gebrauch ‚englisch-klingender‘ verbregierter Präpositionen im Deutschen (Gibson/Hufeisen/Libben 2001: 144). Diese Strategie ordnen Gibson/Hufeisen/Libben (2001: 144) als Hyperkorrektur („over correction“) ein. Diese Interpretation legt nahe, dass linguistisches Wissen aus bereits erlernten Sprachen einen Einfluss auf den Gebrauch von Präpositionen und PP in der Zielsprache Deutsch ausüben kann. Dieser Einfluss kann sich auch in Form von Hyperkorrekturen (=Fehlern) manifestieren.

In einer weiteren Untersuchung können Hufeisen/Gibson (2002) ebenfalls keinen positiven Einfluss der Kenntnisse englischer Präpositionen auf den Gebrauch von Präpositionen im Deutschen feststellen. Hufeisen/Gibson (2002) untersuchen bei insgesamt 21 Lernenden des Deutschen als zweiter Fremdsprache, die allesamt über Kenntnisse der englischen Sprache verfügen, den Gebrauch von lokalen PP mit den Präpositionen *auf*, *an* und *in*. Auch in dieser Untersuchung zeigen die Ergebnisse keinen positiven Einfluss der Englischkenntnisse auf den korrekten Gebrauch deutscher Präpositionen bei sprachäquivalenten PP wie *Das Schloss liegt in England* und *The castle is/lies in England* (Hufeisen/Gibson 2002: 74) im Vergleich zu nicht sprachäquivalenten PP¹³⁹. Die Autorinnen

¹³⁷ Beide Arbeiten entstanden in den 1970er Jahren. Zu dieser Zeit waren Fehleranalysen geprägt von Annahmen der Kontrastivhypothese (vgl. Kuhberg 2001). Somit erklärt sich die dominante Ausrichtung auf den Einfluss der L1.

¹³⁸ Die untersuchten Lernenden beherrschen die deutsche Sprache auf einem fortgeschrittenen Kompetenzniveau (Gibson/Hufeisen/Libben 2001: 141).

¹³⁹ Als Gegenbeispiel zu äquivalenten PP im Deutschen und Englischen vergleichen Gibson/Hufeisen (2002: 74) die englische Konstruktion *to be interested in something* mit der deutschen *sich interessieren für etwas*. An dieser Stelle möchte ich jedoch anmerken, dass die Autorinnen bei ihrer Darstellung ähnlicher bzw. verschiedener PP-Realisierungen im Deutschen und Englischen zwei

schließen aus diesem Ergebnis, dass bei sprachäquivalenten PP die Kenntnisse der englischen Sprache förderlich wirken und somit zur korrekten Präpositionswahl im Deutschen führen; gleichzeitig würden sich bei nicht äquivalenten PP jedoch keine Interferenzerscheinungen zeigen (Hufeisen/Gibson 2002: 79). Damit zeigt sich ein mehrdeutiges Ergebnis in Bezug auf den Einfluss einer bereits erlernten Sprache auf den Gebrauch von Präpositionen in der Zielsprache (Hufeisen/Gibson 2002: 80), das zu einem gewissen Grad von den Ergebnissen zum Gebrauch verbregierter Präpositionen (Gibson/Hufeisen/Libben 2001, s.o.) abweicht.

4.1.3 Sprachkompetenz

Die Forschungsliteratur deutet darauf hin, dass die Schwierigkeiten im Gebrauch von Präpositionen und PP durch Deutschlernende, die sich vorrangig in Fehlern im Präpositionsgebrauch manifestieren, mit einem bestimmten Sprachstand bzw. der Sprachkompetenz¹⁴⁰ in Verbindung stehen. Die Fehlertypen *Wahl* und *Auslassung* gehören zu den typischen Fehlern im Gebrauch von Präpositionen und PP in der Zielsprache Deutsch, s. die folgenden Beispiele aus der Forschungsliteratur:

Der Bus fährt in die Jugendherberge (Grießhaber 2007b: 373).
wenn ich dort, unsere Heimat arbeite (Keim 1984: 140¹⁴¹)

Die Präpositionswahl bzw. die Realisierung der Präposition stellt eine der größten Herausforderungen für Deutschlernende dar (vgl. u. a. Turgay 2010a, Grießhaber 2007b, Hepsöyler/Liebe-Harkort 1991, Apeltauer 1987, Clyne 1968, Keim 1984, Diehl et al. 2000, Hufeisen/Gibson 2002, Hawthorne 2005, Roe 2007). Nur vereinzelt findet man (zusätzlich) Hinweise bzw. Thematisierungen von inkorrekten Hinzufügungen von Präpositionen (Lütke 2008, Kaltenbacher/Klages 2006, HPD 1977, Bouillon 1977 und Roe 2007). Inkorrekte Positionierungen von Präpositionen finden sich bei DaZ-Kindern nicht (vgl. Gutzmann/Turgay 2011, Kuhberg 1990), wohingegen sie im DaZ-Erwerb erwachsener Lernender vereinzelt beobachtet werden können (Meyer-Ingwersen 1975: 76, Keim 1984: 143).

Die einzelnen Fehlertypen im Präpositionsgebrauch (vor allem Auslassung und Wahl) scheinen mit der Sprachkompetenz in Verbindung zu stehen. Apeltauer (1987) führt eine

verschiedene syntaktische Kontexte und damit zwei verschiedene syntaktische Funktionen von PP gegenüberstellen. Bei *Das Schloss liegt in England* bzw. *The castle is/lies in England* liegt eine adverbiale PP vor während beim Paar *to be interested in something* bzw. *sich interessieren für etwas* eine Objekt-PP vorliegt. Damit vergleichen die Autorinnen folglich zum einen die Realisierung von Raumreferenzen und zum anderen die Realisierung von Objekten. Damit wird jedoch implizit der Eindruck erweckt, dass es bei Raumreferenzen Übereinstimmungen in beiden Sprachen gibt und bei Objekten nicht. Dies ist jedoch nicht der Fall, wie bereits bei Gibson/Hufeisen/Libben (2001) deutlich wird. Gibson/Hufeisen (2002: 74) gehen an dieser Stelle nicht näher auf die untersuchten PP-Kontexte ein. Sie sprechen vorrangig von lokalen Präpositionen und „locative phenomena“, die in ihrer Untersuchung im Fokus stehen. Jedoch finden sich in den untersuchten PP-Kontexten sowohl adverbiale PP als auch Objekt-PP, auch wenn letztere einen vergleichsweise sehr geringen Anteil ausmachen (Gibson/Hufeisen 2002: 87f.).

¹⁴⁰ Ich gebrauche hier die Begriffe Sprachstand und Sprachkompetenz synonym.

¹⁴¹ Keim (1984: 140) gibt hier als Rekonstruktion *wenn ich dort in unserer Heimat arbeite* an.

Langzeituntersuchung mit vier Deutschlernenden (L1 Türkisch) über einen Zeitraum von einem Jahr durch und analysiert in diesem Zusammenhang den Gebrauch von Präpositionen. Obwohl er zum Schluss kommt, dass Präpositionen als Indikator für die Ermittlung des Sprachstands von Lernenden „nur bedingt geeignet“ sind (Apeltauer 1987: 230), zeigen seine Daten dennoch Tendenzen, die dafür sprechen, dass Auslassungen von Präpositionen vor allem für eine niedrige Sprachkompetenz charakteristisch sind, während inkorrekt gewählte Präpositionen und Übergeneralisierungen vorrangig auf eine fortgeschrittene Sprachkompetenz deuten (Apeltauer 1987: 222). Zu einem ähnlichen Ergebnis kommt ebenfalls Kuhberg (1990: 26ff.) in seiner longitudinalen Untersuchung zum Gebrauch sprachlicher Strukturen bei Lernenden mit Türkisch und Polnisch als L1. Auch er stellt fest, dass die Auslassung von grammatischen Funktionswörtern – u. a. den Präpositionen – im kindlichen DaZ-Erwerb vor allem charakteristisch ist für einen niedrigen Sprachstand (Kuhberg 1990: 40). Angelehnt an Apeltauer (1987) entwickelt Lütke (2010: 46ff., s. auch Lütke 2011) auf Basis von Daten Deutschlernender mit unterschiedlichen Erstsprachen ein deskriptives Stufenmodell zum Erwerb und Gebrauch von Präpositionen in der Zielsprache Deutsch. Ihre Untersuchungen zeigen eine Beziehung zwischen Fehlertyp und Sprachstand: Der Auslassung von Präpositionen auf dem niedrigsten Sprachstand folgt die Verwendung einer Einheitspräposition und damit verbundene Übergeneralisierungen, die sich in der inkorrekten Wahl von Präpositionen manifestieren, bevor auf einem höheren Sprachstand zielsprachlich korrekte Präpositionen und PP sowie deren Ausbau zu beobachten sind (Lütke 2010: 46 f.). Entsprechende Stadien nennen auch Gutzmann/Turgay (2011: 198): In Bezug auf den DaZ-Erwerb der PP durch Kinder nehmen auch sie auf Basis ihrer Datenauswertungen an, dass die Auslassung von Präpositionen charakteristisch ist für die frühen Stadien im Erwerb des Deutschen und damit für einen niedrigen Sprachstand¹⁴². Pienemann (1981: 50) kommt in seiner Analyse von Daten italienischsprachiger Lernender (Kinder) dahingegen zu einem abweichenden Ergebnis: In seinen Daten ist die Auslassung von Präpositionen auf einem niedrigen Sprachstand nur sehr selten zu beobachten.

In Studien zum DaZ-Erwerb durch Erwachsene lassen sich unterschiedliche Ergebnisse in Bezug auf das Verhältnis von Präpositionsgebrauch und Sprachstand finden: Während Clyne (1968: 138) die Auslassung von Präpositionen durch MigrantInnen in Deutschland noch als ein allgemeines Merkmal „der Gastarbeitersprache“ beschreibt und dabei die unterschiedlichen Sprachkompetenzen seiner ProbandInnen nicht thematisiert¹⁴³, zeigen die Analysen des HPD (1977) differenziertere Befunde. Die Untersuchungen zum DaZ-Erwerb von GastarbeiterInnen mit Spanisch und Italienisch als L1 illustrieren, dass Auslassungen von Präpositionen vor allem

¹⁴² Dabei nennen Gutzmann/Turgay (2011: 198f.) identische Stadien für den PP-Erwerb in Erst- und Zweitsprache, stellen in ihrer Untersuchung jedoch fest, dass die von ihnen analysierten Daten von Kindern mit deutscher Muttersprache fast ausschließlich fortgeschrittene Sprachstände abbilden, und Auslassungen von Präpositionen sind hier nur äußerst selten belegt.

¹⁴³ Clyne (1968: 137) erwähnt beispielsweise, dass einige wenige der untersuchten GastarbeiterInnen, die Präpositionen verwenden, „meistens die falschen“ Präpositionen wählen. Es bleibt jedoch unklar, welchen Sprachstand diese ArbeiterInnen haben. Es wäre durchaus möglich, dass just diejenigen GastarbeiterInnen, die einen vergleichsweise höheren Sprachstand aufweisen, vermehrt inkorrekte Präpositionen wählen.

bei Lernenden mit einem niedrigen Sprachstand zu beobachten sind. Dahingegen sinkt der Anteil dieses Fehlertyps mit Zunahme des Sprachstands (HPD 1977: 102). Interessant ist ebenfalls die Beobachtung, dass Lernende mit einem sehr hohen Sprachstand mehr Präpositionen gebrauchen als deutsche MuttersprachlerInnen (HPD 1977: 102). Des Weiteren beobachten die WissenschaftlerInnen, dass „der Gebrauch der Präpositionen durch die Lernergruppe zum Teil noch von der Zielvarietät ab[weicht]“ (104) – unklar bleibt hier jedoch, wie der Präpositionsgebrauch durch DaZ-Lernende vom zielsprachlichen Gebrauch abweicht, ob also z. B. die Präposition inkorrekt gewählt wird oder ob sie inkorrekt hinzugefügt wird.

Meisel (1977) untersucht in einer Pilotstudie zum ZISA-Projekt (Clahsen/Meisel/Pienemann 1983) mündliche Sprachdaten spanischer und türkischer DaZ-Lernender und beobachtet ebenfalls, dass vor allem Lernende mit niedrigem Sprachstand u. a. Präpositionen häufig auslassen. Meisel (1977: 108) beschreibt die Auslassung von Präpositionen als eine Strategie der Simplifizierung, die vor allem für die frühen Stadien des (ungesteuerten) Spracherwerbs charakteristisch ist, da sie der Sicherung einer effizienten Kommunikation mithilfe minimaler sprachlicher Mittel dient. Übergeneralisierungen, zu denen ebenfalls die inkorrekte Wahl einer Einheitspräposition zählt, stellen im Gegensatz dazu Simplifizierungen dar, die dem Ausbau der Lerner Sprache dienen, da sie das Testen neuer Hypothesen über den Aufbau des zielsprachlichen Systems anzeigen können. Eine ähnliche Beobachtung äußert Meyer-Ingwersen (1975) auf Basis der Analyse schriftlicher Daten von SchülerInnen mit Türkisch als L1: Abweichungen in der Wahl von Präpositionen markieren im Vergleich zu Auslassungen „Fehler, die man erst auf der Basis bestimmter Lernfortschritte machen kann“. Damit signalisieren solche Fehler nach Meyer-Ingwersen (1975: 76) einen Fortschritt im Erwerb des Deutschen und damit einen tendenziell höheren Sprachstand. Auch bei Boullion (1977: 53) findet sich die Annahme, dass wahlbedingte Präpositionsfehler im Verlauf des Spracherwerbs und damit mit Anstieg des Sprachstands zunehmen.

Keim (1984: 215) stellt bei ihrer Analyse mündlicher Daten erwachsener DaZ-Lernender mit Türkisch als L1 und unterschiedlichem Sprachstand des Deutschen ebenfalls fest, dass Auslassungen von Präpositionen charakteristisch zu sein scheinen für einen niedrigen Sprachstand, d. h. weniger entwickelte Lerner Sprachen. Ob die inkorrekte Wahl einer Präposition dahingegen eine fortgeschrittene Sprachkompetenz markiert, wird bei Keim (1984) nicht expliziert.

Wie Hawthorne (2005: 131) in seiner Untersuchung zur Kenntnis lokaler und direktonaler PP durch 30 Lernende mit Englisch als L1 am Beispiel der Präpositionen *an* und *auf* zeigt, finden sich wahlbedingte Präpositionsfehler sowohl auf einem niedrigen als auch auf einem vergleichsweise hohen Sprachstand der Lernenden. Hawthorne (2005) berichtet an keiner Stelle von Auslassungen von Präpositionen. Dieser Umstand ist jedoch vorrangig auf die Datenerhebungsmethode zurückzuführen: Hawthorne (2005) erhebt seine Daten im Rahmen einer ‚forced elicitation‘ task sowie einer ‚quasi-production task‘ (129f.). In beiden Kontexten besteht die Aufgabe darin, aus fest vorgegebenem Sprachmaterial auszuwählen bzw. dieses zu vollständigen Sätzen zu konstruieren. Das präsentierte Sprachmaterial enthält u. a. Präpositionen, und die Lernenden wurden in beiden Aufgaben aufgefordert, das präsentierte Sprachmaterial zu nutzen, sodass eine Auslassung von Präpositionen aufgrund

der Aufgabenkonzipierung erst gar nicht möglich war. Man kann folglich annehmen, dass der Gebrauch bzw. die Realisierung von Präpositionen durch den stark kontrollierten Erhebungskontext beeinflusst wurde¹⁴⁴.

4.1.4 Syntaktische Funktion der PP

Bisherige Untersuchungen zum Gebrauch von Präpositionen und PP lassen weiterhin vermuten, dass die Schwierigkeiten im Präpositionsgebrauch, die durch die verschiedenen Fehlertypen widerspiegelt werden, mit dem syntaktischen Kontext zusammenhängen, in dem Präpositionen und PP produziert werden.

So stellt Kuhberg (1990: 31) am Rande fest, dass der von ihm untersuchte Lerner mit L1 Türkisch Präpositionen in Objekt-PP auch dann noch häufig auslässt (*und wir warten den Lehrer*), wenn die temporal und lokal gebrauchten adverbialen PP bereits zu seinem „festen lernersprachlichen Inventar gehören“. Aus den Auslassungen von Präpositionen und aus der generellen Beobachtung, dass Objekt-PP spärlich in den Daten seiner untersuchten Lernenden belegt sind, leitet Kuhberg (1990: 31) die Annahme ab, dass sich der Erwerb verbregierter Präpositionen (und damit von Objekt-PP) langsamer vollzieht als der Erwerb nicht verbregierter Präpositionen. Aus dieser Schlussfolgerung lässt sich ebenfalls ableiten, dass die Auslassung von Präpositionen als wichtiger Indikator für deren Nicht-Erwerb angesehen wird. Andere Untersuchungen zeigen, dass Auslassungen von Präpositionen nicht ausschließlich in Objekt-PP beobachtet werden. So zeigen z. B. die Auswertungen von Apeltauer (1987), Hepsöyler/Liebe-Harkort (1991), Meyer-Ingwersen (1975) und Diehl et al. (2000), dass dieser Fehlertyp in verschiedenen syntaktischen Kontexten von PP beobachtet wird. Bei Diehl et al. (2000: 317) stellen Auslassungen von Präpositionen jedoch unabhängig von der syntaktischen Funktion der PP „Einzelfälle“ dar.

Hinweise auf einen späten Erwerb von Präpositionen in Objekt-PP liefern auch die Auswertungen von Bouillon (1977), Diehl et al. (2000) und Harnisch (1993). Bei der Analyse der Fehlertypen im Gebrauch deutscher Präpositionen und PP durch DaF-Lernende mit L1 Französisch stellt Bouillon (1977: 52) fest, dass "am Ende des dritten Schuljahres noch fast keine Präpositionen in Präpositionalobjekten" festzustellen sind, obwohl PP in anderen syntaktischen Funktionen bereits von den Lernenden gebraucht werden. Bouillon (1977: 52) führt diesen Umstand auf den sprachlichen Input im Unterricht zurück, wenn er sagt, dass Objekt-PP „anscheinend erst später in den Lehrbüchern eingeführt werden“. Auch Diehl u. a. (2000: 276), die ebenfalls schriftliche Sprachproduktionen DaF-Lernender mit L1 Französisch über einen Zeitraum von drei Jahren auswerten, erwähnen, dass der Gebrauch von Objekt-PP erst in höheren Klassen zu beobachten ist. Harnisch (1993: 321) untersucht Sprachproben von 14 Lernenden mit L1 Türkisch und stellt u.a. ebenfalls fest, dass Verben mit Präpositionsreaktion „nicht oder nur vereinzelt“ in ihren Daten belegt sind. Nach Harnisch (1993: 321) zählen Verben mit Präpositionsreaktion zu den "differenziertere[n] sprachliche[n]

¹⁴⁴ Es ist noch hinzuzufügen, dass die Datenerhebungsmethode Hawthornes (2005) sich aus der zugrundeliegenden linguistischen Theorie, der Generativen Grammatik (GG), bedingt, die eine Analyse von (freien) Sprachproduktionen (Performanzdaten) ablehnt, da diese nach Auffassung der GG die sprachliche Kompetenz von Lernenden nicht adäquat anzeigen können (vgl. Elsen 2014: 127).

Mittel[n]" der deutschen Sprache und werden daher erst in einem späteren Stadium erworben.

Pfaff (1984) und Kaltenbacher/Klages (2006) stellen in ihren Untersuchungen mit Lernenden unterschiedlicher Erstsprachen fest, dass Hinzufügungen von Präpositionen in Kontexten direkter Objekte produziert werden, um diese formal zu markieren. Häufig werden hier die Präpositionen *bei*, *zu* (*der Mädchen gibt den Telefon zu seine Freund*, Pfaff 1984: 294) oder *für* (*der Pferd schenkt für eine Hund ein Keks*, Kaltenbacher/Klages 2006: 87) gebraucht. Kaltenbacher/Klages (2006: 87) schließen aus diesen Beispielen, dass „den Kindern die zielsprachliche Form für eine von ihnen erkannte Funktion" noch fehlt und sie aus diesem Grund Präpositionen bzw. PP zur Markierung von indirekten Objekten verwenden¹⁴⁵. Analysen des HPD (1977: 108b) zeigen eine ähnliche Strategie in Bezug auf das Hinzufügen von Präpositionen: Hier ist jedoch festzustellen, dass Präpositionen tendenziell eher in Kontexten von direkten Objekten hinzugefügt werden. Allerdings ist das (an dieser Stelle einzige, ohne weiteren Kontext) aufgeführte Beispiel *spricht in Deutsch* (HPD 1977: 108b) nicht eindeutig interpretierbar – eine Interpretation der PP *in Deutsch* als direktes Objekt ist möglich, es könnte sich jedoch ebenso um die inkorrekte Wahl der Präposition *in* handeln, wenn man sich als Zieläußerung *ich spreche auf Deutsch* vorstellt (vgl. dazu Helbig/Buscha 2001b: 365).

Die Markierung indirekter Objekte mittels Präpositionen bzw. PP beobachten auch Pepouna et al. (2018: 76) bei erwachsenen DaF-Lernenden mit Bantu und Semibantu als Erstsprachen¹⁴⁶. Bantu zeichnet sich dadurch aus, dass indirekte Objekte lediglich durch Wortstellung markiert werden während Semibantu das indirekte Objekt mittels einer PP kodiert. Interessant ist nun, dass im Gebrauch der Zielsprache Deutsch hinzugefügte Präpositionen in Kontexten von Dativobjekten (z. B. *Die Mutter gibt das Buch zu dem Kind*) bei Lernenden beider Erstsprachen zu beobachten sind, wenn auch in einem geringen Ausmaß (Pepouna et al. 2018: 76). Leider berücksichtigen die AutorInnen diesen Fehlertyp nicht für ihre Analysen. Dementsprechend gehen sie auf diese Verwendungsweise von Präpositionen nicht näher ein.

Einige Untersuchungen zum Gebrauch von Präpositionen und PP durch Deutschlernende unterscheiden bei der Datenauswertung nicht zwischen den syntaktischen Kontexten bzw. Funktionen der entsprechenden PP, wenn die syntaktischen Funktionen denn überhaupt explizit thematisiert werden. Hierzu zählen vor allem die Untersuchungen von Pfaff (1984), Apeltauer (1987), Pienemann (1981), Meyer-Ingwersen (1975) und Clyne (1968). In diesen Untersuchungen lassen sich die syntaktischen Kontexte des Gebrauchs von Präpositionen lediglich anhand der angegebenen Datenbeispiele ableiten¹⁴⁷. Das HPD (1977:

¹⁴⁵ Derartige Strukturen finden sich auch in der Zielsprache wieder – so verfügt das Deutsche z. B. über Verben mit Dativobjekten und parallelen präpositionalen Varianten (*schreiben* + Dat. und *schreiben* + an-PP, Eisenberg 2013: 265.). Zur Bedeutung solcher präpositionalen Varianten s. z. B. Proost (2015).

¹⁴⁶ Die Sprachen gehören zu den heimischen Sprachen im Kamerun (Afrika) (Pepouna et al 2018: 73).

¹⁴⁷ Beispiele aus Pfaff (1984): *Ich will Türkei gehen/Die liegt da oben in das Tisch* (274), aber auch *gibt den Ball zu den Jungs* (294) oder *ich war in Neu-Istanbul gegangen/und auf einmal klopft es in der Tür* (296). Beispiele aus Apeltauer (1987): *Die schneid in die Tisch/Der Vater spielt Kind/Fliegt die Kopf*

101) unterscheidet zumindest zwischen attributiven PP und „andere[n] Präpositionalphrasen“, eine weitere Differenzierung der nicht-attributiven PP erfolgt jedoch nicht. Roe (2007) analysiert valenzbedingte Schwierigkeiten von Deutschlernenden mit L1 Englisch. Bei ihm stehen PP-Ergänzungen im Vordergrund, eine Differenzierung nach Objekt-PP und adverbialer PP wird aber nicht vorgenommen. Andere Untersuchungen fokussieren explizit ausgewählte syntaktische Kontexte von PP: In der Regel handelt es sich hierbei jedoch um adverbiale PP (Grießhaber 2007b, Becker 2006, Hepsöyler/Liebe-Harkort 1991, Turgay 2010a, b, 2011, Gutzmann/Turgay 2011, Lütke 2008, Bryant 2012, Carroll/Becker 1993, Becker 2012, Hufeisen/Gibson 2002, Hawthorne 2005). Die einzig mir bekannte Untersuchung, die sich ausschließlich Präpositionen in Objekt-PP widmet ist jene von Gibson/Hufeisen/Libben (2001), wobei sie dort ausschließlich die Kenntnis der verbregierten Präpositionen in den Fokus nehmen.

In den erwähnten Untersuchungen wird deutlich, dass Präpositionen und PP in der Regel in ihren prototypischen Verwendungen als adverbiale PP zum Ausdruck von Lokalität oder Direktionalität untersucht werden. Dieser Schwerpunkt ist natürlich eng mit dem jeweiligen Forschungsinteresse einzelner Studien verknüpft. Lütke (2008), Bryant (2012), Carroll/Becker (1993) und Becker (2012) fokussieren beispielsweise den Erwerb und die sprachliche Realisierung von Raumreferenzen bzw. Lokalisierungsausdrücken. Somit stehen lokale Verwendungsweisen von PP im Vordergrund und die Datenerhebungsmethoden sind darauf abgestimmt. Andere Einschränkungen bei den syntaktischen Funktionen von PP können aufgrund von Datenmangel erklärt werden: So räumt Turgay (2010a: 5) z. B. in ihrer Untersuchung ein, dass sie nicht zwischen unterschiedlichen syntaktischen Funktionen von PP unterscheidet, da ihre Datengrundlage „fast ausschließlich“ lokale adverbiale PP enthält. Eine Vernachlässigung der anderen syntaktischen Funktionen (wie z. B. Objekt-PP) ist somit gerechtfertigt¹⁴⁸.

Zu den Untersuchungen von Harnisch (1993) und Kaltenbacher/Klages (2006) ist noch hinzuzufügen, dass diese zwar vereinzelt auf Verben mit Präpositionsreaktion eingehen, jedoch nicht im Rahmen ihrer Auswertung des Präpositionsgebrauchs. Harnisch (1993: 320) differenziert bei der Auswertung des Gebrauchs von Präpositionen nicht nach den syntaktischen Kontexten von PP. Daraus lässt sich indirekt ableiten, dass Harnisch bei der Auswertung des Präpositionsgebrauchs lediglich solche Präpositionen betrachtet, die nicht Bestandteil von Objekt-PP sind. Ein expliziter Hinweis zur syntaktischen Funktion der PP fehlt jedoch an entsprechender Stelle. Ein ähnlicher Fall liegt bei Kaltenbacher/Klages (2006) vor: Auch sie konzentrieren sich bei der Auswertung des Gebrauchs von Präpositionen und PP auf „räumliche Präpositionen“ (d. h. lokale und direktive adverbiale PP) (88ff.).

Untersuchungen, die unterschiedliche syntaktische Kontexte von PP erwähnen, beziehen sich jedoch in keinem der Fälle systematisch auf ausgewählte Präpositionen. D. h.

(222). Hier wird deutlich, dass unterschiedliche syntaktische Kontexte von PP berücksichtigt, aber nicht explizit differenziert werden.

¹⁴⁸ Die Auswertungen in Turgay (2010b, 2011) und Gutzmann/Turgay (2011) beruhen auf der Datengrundlage aus Turgay (2010a). Somit liegt auch dort eine Einschränkung auf lokale adverbiale PP vor.

der Gebrauch von Präpositionen und PP in unterschiedlichen syntaktischen Kontexten wird, wenn überhaupt, vereinzelt thematisiert, jedoch keinem systematischen Vergleich unterzogen (im Ansatz stellt Kuhberg 1990 hier eine Ausnahme dar, s.o.).

4.1.5 Semantischer Gehalt der Präposition

Wenige Untersuchungen im DaZ-/DaF-Kontext widmen sich dem Einfluss des semantischen Gehalts einer Präposition auf deren Gebrauch. In einigen Untersuchungen findet man jedoch Hinweise auf einen möglichen Einfluss der Bedeutung einer Präposition bzw. des semantischen Gehalts einer Präposition.

Turgay (2010a) erwähnt einen Einfluss der semantischen Komplexität einer Präposition auf ihren Gebrauch. In ihrer Untersuchung von DaZ-Lernenden (Kinder) stellt sie (Turgay 2010a: 282) fest, dass „semantisch komplexere[-] Präpositionen“ wie z. B. die kausale Präposition *wegen*, unabhängig von der L1 später erworben werden als semantisch einfache Präpositionen (wie z. B. *in*, *auf*, *bei*). Was hier genau unter ‚semantisch komplex‘ verstanden wird, bleibt leider unklar. Jedoch kann vermutet werden, dass hier die ‚Konkretheit‘ der semantischen Relation einer Präposition gemeint ist: Lokale (konkrete) Relationen (*in*, *auf*, *bei*) werden früher erworben als z. B. kausale (abstrakte) Relationen (*wegen*). In beiden Fällen wird der Präposition ein semantischer Gehalt zugeschrieben, d. h. die Präposition wird als semantisiert angesehen. Turgays (2010a) Ergebnisse zeigen demnach, dass innerhalb von semantisierten Präpositionen Unterschiede im Erwerb und Gebrauch der Präpositionen beobachtet werden können. Die Ergebnisse von Hufeisen/Gibson (2002) deuten ebenfalls auf solch eine Interpretation hin: In den Ergebnissen ihrer Untersuchung zum Gebrauch der Präpositionen *auf*, *an* und *in* erkennen Hufeisen/Gibson (2002) vereinzelt Hinweise darauf, dass eine desemantisierte Präposition (z. B. in Objekt-PP wie *Der Hund starb an Krebs*)¹⁴⁹ mehr Lernschwierigkeiten, d. h. mehr Abweichungen bei der Präpositionswahl aufweist als dieselbe Präposition im semantisierten Gebrauch (z. B. *Das Schild hängt an der Tür*) (vgl. Hufeisen/Gibson 2002: 76). Die Ergebnisse spiegeln jedoch lediglich Tendenzen wider, da die Datenmenge laut den Autorinnen zu gering ist, um aussagekräftige Schlüsse ziehen zu können (Hufeisen/Gibson 2002: 80). Die Ergebnisse von Baten (2009) zeigen Parallelen zur Beobachtung von Hufeisen/Gibson (2002). In einer Umfrage mit 18 DaF-Lehrkräften, die Baten (2009: 99) durchführte, wird der Gebrauch von desemantisierten Präpositionen (wie sie z. B. in Objekt-PP vorkommen) als Problembereich bei Deutschlernenden wahrgenommen. Auch Balcı (2006: 240) erwähnt, dass vor allem Präpositionen in Objekt-PP „ein[en] Alptraum für Deutschlernende“ darstellen. Balcı (2006: 240) führt dies auf den geringen semantischen Gehalt zurück, den Präpositionen in Objekt-PP üblicherweise aufweisen. Jedoch geht er in seiner Untersuchung nicht näher auf den Gebrauch bzw. Gebrauchsschwierigkeiten solcher Präpositionen ein.

¹⁴⁹ Obwohl Hufeisen/Gibson (2002) an dieser Stelle lediglich das Beispiel einer Objekt-PP nennen, zählen sie den temporalen Gebrauch einer Präposition (in einer adverbialen PP) ebenfalls zum Beispiel einer Präposition mit geringem semantischem Gehalt (Hufeisen/Gibson 2002: 76). Bei der Beschreibung der Ergebnisse jedoch merken die Autorinnen an, dass im Datenmaterial lediglich eine Instanz des temporalen Gebrauchs einer Präposition bzw. PP vorliegt (Hufeisen/Gibson 2002: 80).

In ihren Untersuchungen zum Erwerb von Raumreferenzen durch erwachsene DaZ-Lernende finden sich bei Caroll/Becker (1993: 147) Hinweise darauf, dass lexikalische Formen (z. B. Präpositionen) zunächst in prototypischen Kontexten erworben und verwendet werden, d. h. prototypische Form-Funktionsverbindungen werden zuerst erlernt. Überträgt man diesen Befund auf die Form-Funktionsverbindungen von Präpositionen und PP, müssten sich demnach in Abhängigkeit der Prototypikalität der Form-Funktionsverbindungen Unterschiede im Gebrauch entsprechender Präpositionen bzw. PP zeigen. Einen Hinweis darauf findet man im Ansatz bereits bei Viorel (1973). Viorel (1973: 355) führt Fehler im Gebrauch von deutschen Präpositionen u.a. darauf zurück, dass die Bedeutung der Präposition in bestimmten Kontexten von der ursprünglichen Bedeutung dieser Präposition abweicht. Als Beispiel dient die Verwendung der Präposition *mit* in temporaler Gebrauchsweise (*mit fünf Jahren*) im Vergleich zu ihrer prototypischen Verwendung in Kontexten von ‚Gemeinsamkeit‘ wie z. B. in *mit jmdm. spielen* (vgl. Viorel 1973: 355). In zielsprachlichen Kontexten, in denen die Präposition *mit* in ihrer „Hauptbedeutung“ gebraucht wird, stellt Viorel (1973: 355) weniger wahlbedingte Fehler im Präpositionsgebrauch fest. Viorel (1973) betrachtet den Gebrauch deutscher Präpositionen jedoch aus einer kontrastiven Perspektive und verweist auf den Einfluss der L1 (hier: Rumänisch) sowie die (Nicht-)Entsprechungen rumänischer und deutscher Präpositionsverwendungen und sich daraus ergebenden Schwierigkeiten. Insofern ist unklar, inwiefern Viorels (1973) Ergebnisse eindeutig auf den Einfluss von Form-Funktionsverbindungen und den damit einhergehenden unterschiedlichen semantischen Gehalt deutscher Präpositionen schließen lassen.

An dieser Stelle möchte ich festhalten, dass keine mir bekannten Untersuchungen vorliegen, die den semantischen Gehalt von Präpositionen und deren Einfluss auf den Gebrauch von Präpositionen und PP gezielt untersuchen. Lediglich die Untersuchung von Hufeisen/Gibson (2002) präsentiert einen möglichen Ansatz. Diese Untersuchung deutet auf einen möglichen Einfluss des semantischen Gehalts von Präpositionen auf mögliche Gebrauchsschwierigkeiten im Gebrauch von (desemantisierten) Präpositionen hin. Systematische Untersuchungen zum Einfluss des semantischen Gehalts einer Präposition auf ihren Gebrauch durch Deutschlernende fehlen jedoch meines Wissens gänzlich.

4.1.6 Zusammenfassung

Im Folgenden werden die zentralen Ergebnisse der ausgewählten Einflussvariablen auf den Gebrauch von Präpositionen und PP durch Deutschlernende tabellarisch zusammengefasst (Tab. 17).

Einflussvariable	Erkenntnisse zum Gebrauch von Präpositionen und PP in der Zielsprache Deutsch
Erstsprache (L1) der Lernenden	<ul style="list-style-type: none"> • Gebrauch von Präpositionen = Lernschwierigkeit, unabhängig von der L1 • Einfluss der L1 auf Erwerbstempo und Fehlertyp im Präpositionsgebrauch: Wahlbedingte Fehler bei typologischer Nähe der L1; Inkorrekte Position von Präpositionen nur in Einzelfällen. • In bisherigen Forschungen im Fokus: Lernende mit L1 Türkisch; DaZ-Erwerb
Sprachkompetenz	<ul style="list-style-type: none"> • Zusammenhang von Gebrauchsschwierigkeiten (d. h. Fehlertypen) und Sprachkompetenz. Tendenzen: Auslassungen von Präpositionen vorrangig bei niedriger Sprachkompetenz; wahlbedingte Fehler bei höherer Sprachkompetenz.
Syntaktische Funktion der PP	Bisher keine empirisch fundierten Erkenntnisse. Annahmen: <ul style="list-style-type: none"> • Erwerb (und Gebrauch) adverbialer PP vor Objekt-PP • Fehlertypen: Auslassungen in Objekt-PP; Hinzufügungen in Kontexten von direkten/indirekten Objekten
Semantischer Gehalt der Präposition	Bisher keine empirisch fundierten Erkenntnisse. Annahme: <ul style="list-style-type: none"> • Mehr Gebrauchsschwierigkeiten bei desemantisierten Präpositionen im Vergleich zu semantisierten Präpositionen.

Tabelle 17: Forschungsstand zu ausgewählten Einflussvariablen im Gebrauch von Präpositionen und PP in der Zielsprache Deutsch: Zusammenfassung.

Die Variablen *Erstsprache (L1) der Lernenden* und die *Sprachkompetenz* können als gut erforschte Variablen gelten, da hierzu empirisch fundierte Erkenntnisse vorliegen. Zu den Variablen *Syntaktische Funktion der PP* und *Semantischer Gehalt der Präposition* liegen bisher keine systematischen Untersuchungen vor, es lassen sich jedoch jeweils Annahmen aus bisherigen Forschungsarbeiten ableiten.

4.2 Präpositionen und PP in der (deutschen) Lernerkorpusforschung

4.2.1 Einleitung

In diesem Kapitel gebe ich einen kurzen Überblick über den Forschungsstand zum Untersuchungsgegenstand Präpositionen und PP in der Lernerkorpusforschung. Im Fokus stehen dabei Arbeiten, die primär an den Schnittstellen zur Zweit- und Fremdspracherwerbsforschung (Z/FSE) (4.2.2) sowie zur maschinellen Sprachverarbeitung (NLP) bzw. automatischen Erkennung von Fehlern im Präpositionsgebrauch (4.2.3) zu verorten

sind. Dabei sei bereits an dieser Stelle angemerkt, dass für die Zielsprache Deutsch an der Schnittstelle zu NLP (bisher) keine mir bekannten Arbeiten vorliegen. Aus diesem Grund ziehe ich hier Arbeiten zur Zielsprache Englisch heran. Diese Arbeiten zeigen, inwiefern Daten aus Lernerkorpora für die Entwicklung automatischer Systeme eingesetzt werden können bzw. welchen Mehrwert linguistische Untersuchungen in Lernerkorpora für die Entwicklung eben solcher Systeme bieten – beide Aspekte haben eine direkte Relevanz für mögliche zukünftige Ansätze zur Entwicklung automatischer Fehlererkennungssysteme für die Zielsprache Deutsch.

4.2.2 Untersuchungen zum Gebrauch von Präpositionen und PP in deutschen Lernerkorpora

In Bezug auf den Erwerb und Gebrauch von Präpositionen und Präpositionalphrasen (PP) im Kontext DaZ/DaF lässt sich die Anzahl relevanter Untersuchungen in Lernerkorpora an einer Hand abzählen. Dies liegt vor allem daran, dass trotz positiver Entwicklungen durch neue Lernerkorpusprojekte (z. B. Wisniewski 2018a oder Hirschmann/Nolda 2019) im Bereich der deutschen Lernerkorpuslandschaft derzeit immer noch vergleichsweise wenige Lernerkorpora mit Deutsch als Zielsprache verfügbar sind. Die LCW-Liste¹⁵⁰ listet 15 Lernerkorpora mit Zielsprache Deutsch. Dagegen finden sich in der Liste aktuell 102 Lernerkorpora mit Englisch als Zielsprache, somit dominiert das Englische derzeit immer noch die Lernerkorpuslandschaft (s. auch Granger 2008: 262f.). In der LCB¹⁵¹ finden sich keine Publikationen zum Untersuchungsgegenstand Präpositionen bzw. PP im DaZ/DaF-Bereich. Auf der Publikationsliste des Lernerkorpus-Projekts Falko¹⁵² der Humboldt Universität zu Berlin finden sich vorrangig Publikationen zur Wortstellung oder Konnektoren. Die bisher wenigen Untersuchungen im MERLIN-Lernerkorpus¹⁵³, das die Datengrundlage dieser Dissertation bildet, behandeln bisher vorwiegend Phänomene, die in den einzelnen Sprachkompetenzstufen bzw. -niveaus des GeR zentral sind (z. B. Wortschatz)¹⁵⁴. Neben meiner Masterarbeit (Weber 2014) sind mir aktuell (Stand Mai 2019) lediglich die lernerkorpuslinguistischen Arbeiten von Breindl (2013) und Sieradz/Bordag (2014) bekannt,

¹⁵⁰ LCW steht für *Learner corpora around the world*, siehe <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html> (Mai 2019).

¹⁵¹ LCB steht für *Learner Corpus Bibliography*, verfügbar unter <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpus-bibliography.html> sowie in einer erweiterten Version über das Literaturverwaltungsprogramm Zotero, zugänglich für MitgliederInnen der Learner Corpus Association. Auch in der erweiterten Version der Bibliographie finden sich (neben meiner eigenen Masterarbeit) keine Publikationen zum Untersuchungsgegenstand Präpositionen oder PP im DaZ/DaF-Bereich.

¹⁵² [https://www.linguistik.hu-](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/publikationen)

[berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/publikationen](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/publikationen) (Mai 2019).

¹⁵³ Siehe https://merlin-platform.eu/C_download.php (Mai 2019).

¹⁵⁴ Da MERLIN primär zur Illustrierung und Validierung der GeR-Niveaus aufgebaut wurde (vgl. Abel et al. 2014, s. auch Kap. 6) und damit eng mit der Sprachtestforschung verbunden ist, ist dieser Schwerpunkt in bisherigen Publikationen nachvollziehbar.

die sich dem Gebrauch von Präpositionen durch Deutschlernende widmen¹⁵⁵. Diese Arbeiten zeigen erste Ansätze zur Analyse des Präpositionsgebrauchs durch DaF-Lernende unter Rückgriff auf Lernerkorpusdaten. Die beiden letztgenannten Arbeiten fokussieren den Gebrauch von Verschmelzungen aus Präposition und Artikel. Obwohl die vorliegende Dissertation keinen Schwerpunkt auf den Gebrauch von Verschmelzungen legt, ist die Darstellung der korpuslinguistischen Arbeiten von Breindl (2013) und Sieradz/Bordag (2014) relevant, da sie zu den ersten Arbeiten zählen, die den Gebrauch von Präpositionen auf Basis von Lernerkorpusdaten untersuchen. Im Folgenden gehe ich näher auf die genannten Untersuchungen ein, wobei bei Breindl (2013) und Sieradz/Bordag (2014) primär die eingesetzten Methoden im Fokus stehen. Breindl (2013) und Sieradz/Bordag (2014) untersuchen den Gebrauch von Verschmelzungen aus Präposition und Artikel (*zum, am, im*) sowie deren analytische Vollformen (*zu dem, an dem, in dem*). Beide Untersuchungen greifen primär auf Daten aus dem Lernerkorpus *Falko* (Lüdeling et al. 2008) zurück¹⁵⁶ und konzentrieren sich auf die Abweichungen im Gebrauch von Verschmelzungen und Vollformen.

Breindl (2013) untersucht den Gebrauch der frequenten Verschmelzungsformen *am, im, zum, beim, vom, zur* und *ins* sowie deren Vollformen (*an dem, in dem, etc.*). Unter Rückgriff auf die im Lernerkorpus *Falko* enthaltenen Editier-Tags (s. Kap. 3.3.2) führt Breindl (2013) eine computerunterstützte Fehleranalyse der oben genannten Verschmelzungen und Vollformen durch, wobei ihre Auswertungen vorwiegend rein deskriptiv bleiben. Breindl (2013) betont, dass sie die Untersuchung im Lernerkorpus als Pilotstudie ansieht, deren Ergebnisse Impulse für weitere, hypothesengeleitete empirische Untersuchungen liefern können. Aus diesem Grund finden bei Breindl (2013) die in *Falko* enthaltenen Metadaten (z. B. Erstsprache der Lernenden) bei der Auswertung keine Berücksichtigung. Ebenfalls erfolgt keine statistische Auswertung der Lernerkorpusdaten.

Sieradz/Bordag (2014) widmen sich ebenfalls dem Untersuchungsgegenstand Verschmelzungen aus Präposition und Artikel und deren analytischen Vollformen, wobei sie primär das Ziel haben „die häufigsten Fehlerquellen“ (Sieradz/Bordag 2014: 12) im lernersprachlichen Gebrauch von Verschmelzungen zu ermitteln. Sieradz/Bordag (2014) analysieren Lernerdaten aus dem Lernerkorpus *Falko* und ziehen zusätzlich vergleichende Daten aus einer Fragebogenstudie zur Verwendung von Vollformen und Verschmelzungen mit deutschen MuttersprachlerInnen und Deutschlernenden hinzu. Sieradz/Bordag (2014) führen eine Fehleranalyse der inkorrekt gebrauchten Verschmelzungen und Vollformen im Lernerkorpus *Falko* durch. Die Fehleranalyse erfolgt dabei (scheinbar) ohne Rückgriff auf die im Lernerkorpus enthaltenen Annotationen bzw. Editier-Tags (Sieradz/Bordag 2014: 13f.). Ähnlich zu Breindl (2013) bleiben die Korpusauswertungen bei Sieradz/Bordag (2014) deskriptiv (d. h. nicht hypothesenüberprüfend), Metadaten werden nicht einbezogen und es

¹⁵⁵ Studien zum Gebrauch von Präpositionen und PP auf Basis digitaler Korpora mit Daten von SchülerInnen mit Muttersprache Deutsch liegen z. B. von Glaznieks/Abel (2017) und Glück/Leonardi (i. Dr.) vor.

¹⁵⁶ Breindl (2013) analysiert ebenfalls Daten aus dem Lernerkorpus des Netzwerks Kobalt-DaF (Zinsmeister et al. 2012). Sieradz/Bordag (2014) ziehen ebenfalls Daten von MuttersprachlerInnen hinzu und vergleichen diese u.a. mit Daten im Lernerkorpus *Falko*.

finden sich ebenfalls keine statistischen Auswertungen. Da Sieradz/Bordag (2014) in ihrer Untersuchung auf weitere Daten aus einer Fragebogenstudie zurückgreifen, scheint ihre Untersuchung im Lernerkorpus Falko lediglich ein Mittel zu sein, um sich einen ersten Überblick über potentielle Gebrauchsschwierigkeiten zu verschaffen, den sie im Anschluss „auf Basis einer größeren Datenbasis“ zu differenzieren versuchen (Sieradz/Bordag 2014: 14).

Meine Masterarbeit (Weber 2014, 2015) beschäftigt sich mit dem Gebrauch von Objekten ausgewählter Verbgruppen durch DaF-Lernende. Als Datengrundlage dient das Lernerkorpus MERLIN (damals noch im Aufbau befindend). Weber (2014) vergleicht den Gebrauch ausgewählter Verben mit Kasus-Objekten (z. B. *lieben, helfen*) mit Verben, die eine Objekt-PP regieren (z. B. *warten auf, denken an*); der Fokus liegt auf lernersprachlichen Abweichungen. In Bezug auf den Gebrauch von Objekt-PP zeigt die Untersuchung im Lernerkorpus, dass der Gebrauch der Präpositionen Lernenden Schwierigkeiten bereitet, wobei die Auslassung der Präposition den häufigsten Fehlertyp darstellt (Weber 2014: 49). Eine Analyse des Gebrauchs von Objekt-PP in den einzelnen Kompetenzniveaus zeigt jedoch die Tendenz, dass der Fehlertyp Auslassung vor allem für die niedrigen Kompetenzniveaus charakteristisch ist (Weber 2014: 50f.). Zudem enthalten die analysierten Lernerdaten Hinweise darauf, dass sich je nach Präposition Unterschiede in der Korrektheit und den Fehlertypen im Präpositionsgebrauch zeigen (Weber 2014: 52). Neben dem Kompetenzniveau werden bei Weber (2014) keine weiteren Metadaten für die Untersuchung berücksichtigt.

4.2.3 Arbeiten zur automatischen Erkennung fehlerhafter Präpositionen

An der Schnittstelle der Lernerkorpusforschung zur maschinellen Sprachverarbeitung (NLP) findet eine Beschäftigung mit Präpositionen und PP vermehrt statt im Rahmen maschineller Lernverfahren zur Entwicklung von Systemen zur automatischen Fehlererkennung in schriftlichen Texten. Ich möchte exemplarisch auf ausgewählte Arbeiten in diesem Bereich eingehen. Dabei stehen nicht die technischen Aspekte der Systeme im Vordergrund, sondern inwiefern Daten aus Lernerkorpora in der Entwicklung solcher automatischer Systeme sinnvoll eingesetzt werden können¹⁵⁷. Zu methodologischen Aspekten automatischer Fehlerannotationen s. z. B. Tetreault/Chodorow (2008a¹⁵⁸) und Gamon et al. (2013).

Arbeiten von z. B. Chodorow/Tetreault/Han (2007), Tetreault/Chodorow (2008a, b), Cahill et al. (2013) sowie Gamon et al. (2013) widmen sich der automatischen Fehlererkennung im lernersprachlichen Präpositionsgebrauch und testen bzw. evaluieren mögliche Lernverfahren in diesem Bereich. In den Arbeiten zur automatischen Fehlererkennung im Präpositionsgebrauch dominiert die englische Sprache, entsprechende Arbeiten für das Deutsche sind mir nicht bekannt. Die Arbeiten zur automatischen Fehlererkennung im Bereich Präpositionsgebrauch konzentrieren sich auf die automatische Erkennung ausgewählter Fehlertypen im Präpositionsgebrauch, meist steht dabei die

¹⁵⁷ Für einen umfassenden Überblick zu Ansätzen der automatischen Fehlererkennung in lernersprachlichen Produktionen (Bereich Grammatik und Orthographie) s. Leacock et al. (2014).

¹⁵⁸ Tetreault/Chodorow (2008a) diskutieren die Zuverlässigkeit von Datenannotationen durch einen oder mehrere menschliche RaterInnen zur Evaluation automatischer Annotationsverfahren. Zudem präsentieren sie einen alternativen Annotationsansatz, der eine zeit- und kostengünstige Alternative darstellt.

inkorrekte Wahl einer Präposition im Vordergrund. Chodorow/Tetreault/Han (2007) sowie Tetreault/Chodorow (2008b) berücksichtigen zusätzlich inkorrekt hinzugefügte Präpositionen. Automatische Werkzeuge, die ausgelassene Präpositionen detektieren, liegen meines Wissens nicht vor. Somit decken die bisherigen Ansätze zur automatischen Fehlererkennung nicht das volle Spektrum möglicher frequenter Fehlertypen im Bereich Präpositionsgebrauch ab.

Als Trainingsdaten, die als Grundlage für die Entwicklung von Systemen zur automatischen Fehlererkennung dienen, werden primär linguistische Korpora der Zielsprache (hier: Englisch) genutzt; die automatische Erkennung von Fehlern im Präpositionsgebrauch wird somit primär auf Grundlage korrekter Präpositionsverwendungen trainiert (z. B. Chodorow/Tetreault/Han 2007). Die auf diese Weise trainierten Systeme werden im Anschluss auf weiteren, in der Regel von den Trainingsdaten abweichenden, Daten evaluiert, z. B. an authentischen Lerneräußerungen oder an weiteren muttersprachlichen Daten (Chodorow/Tetreault/Han 2007, Tetreault/Chodorow 2008b). Hier zeigt sich jedoch eine Herausforderung in der Anwendung der entwickelten Systeme: Da die Trainingsdaten von den Evaluationsdaten abweichen, ist die Leistung des entwickelten Systems beeinträchtigt: Das System, das auf Basis korrekter Präpositionsverwendungen entwickelt wird, zeigt erhebliche Probleme im Umgang mit fehlerhaften Präpositionsverwendungen (Chodorow/Tetreault/Han 2007: 28). Einen anderen Ansatz wählen Cahill et al. (2013), die für ihr System zur automatischen Erkennung falsch gewählter Präpositionen auf die Revisionsgeschichten der Artikelseiten in der Wikipedia zurückgreifen und diese Daten als Trainingsdaten für ihr maschinelles Lernverfahren nutzen. Damit beziehen sie für die Entwicklung ihres Systems zur automatischen Erkennung von Präpositionsfehlern sowohl den fehlerhaften als auch zielsprachlichen Gebrauch von Präpositionen mit ein.

Daten aus Lernerkorpora können ebenfalls für die Entwicklung automatischer Fehlererkennung und -korrektur gewinnbringend sein. Gamon et al. (2013) präsentieren drei Ansätze für den Einsatz von Lernerkorpora: 1) zur Ausrichtung der Systementwicklung, 2) als Trainingsdaten und 3) zur Evaluation entwickelter Systeme. Im Folgenden gehe ich im Einzelnen auf diese Ansätze ein.

1) Durch eine Analyse annotierter Lernerkorpora können zentrale Fehlertypen bzw. -kategorien und deren Distribution in Lernerdaten identifiziert werden (128ff.), um somit z. B. wiederum die Ausrichtung und Entwicklung automatischer Werkzeuge maßgeblich zu leiten. Bereits Tetreault/Chodorow (2008b: 871) machten darauf aufmerksam, dass bei der Entwicklung von Systemen zur automatischen Erkennung (und Korrektur) von Fehlern im Präpositionsgebrauch die vorangestellte Analyse typischer lernersprachlicher Abweichungen in diesem Bereich oft vernachlässigt wird. So stellen Tetreault/Chodorow (2008b: 871) in ihrer Analyse von Präpositionsfehlern in Lernertexten (Zielsprache Englisch) fest, dass Präpositionen wie *in*, *to* und *of* am häufigsten inkorrekt verwendet werden und dass die zehn häufigsten Präpositionen 93,8% aller Präpositionsfehler ausmachen. Dies sind nützliche

Informationen, die in die Entwicklung von Systemen zur automatischen Fehlererkennung einfließen müssen¹⁵⁹.

2) Daten aus Lernerkorpora können als Trainingsdaten für die Entwicklung von Systemen zur automatischen Fehlererkennung und -korrektur genutzt werden. Auf diese Weise könnten die entwickelten Verfahren bestmöglich auf ihren späteren Anwendungsbereich – Lernerdaten mit zielsprachlichem und nicht-zielsprachlichem Gebrauch von Präpositionen – angepasst werden. Dieser Ansatz verspricht auch bessere Ergebnisse in der Anwendung automatischer Systeme als der bisher gängige Ansatz, bei dem als Trainingsdaten ausschließlich linguistische Korpora der Zielsprache (in der Regel bestehend aus redigierten Zeitungstexten) eingesetzt werden (s. z. B. Chodorow/Tetreault/Han 2007). Ein ähnlicher Ansatz besteht darin, annotierte Lernerkorpora bzw. die Verteilung von Fehlertypen in Lernerkorpora als eine Vorlage zu nutzen, um systematisch ‚künstliche‘ Fehler in sonst fehlerfreie Daten zu integrieren. Die auf diese Weise entstandenen künstlichen Fehlerdaten („artificial error data“, Gamon et al. 2013: 140) können im Anschluss als Trainingsdaten genutzt werden.

3) Daten aus Lernerkorpora können zur Evaluation der entwickelten Systeme zur automatischen Fehlererkennung und -korrektur eingesetzt werden, so wie dies z. T. bereits erfolgt (vgl. z. B. Chodorow/Tetreault/Han 2007, Tetreault/Chodorow 2008b). Voraussetzung für die Evaluation automatischer Fehlererkennungssysteme sind Lernerkorpora, die bereits fehlerannotiert sind. Eine vielversprechende Alternative zur sehr ressourcenintensiven, manuellen Erstellung von Fehlerannotationen sehen Gamon et al. (2013: 144ff.) in Ansätzen des Crowdsourcings¹⁶⁰. Tetreault/Filatova/Chodorow (2010) und Madnani et al. (2011) illustrieren, wie durch Formen des Crowdsourcings (schnell und kostengünstig) Fehlerannotationen gewonnen und evaluiert werden können. Beide Arbeiten berichten von vielversprechenden ersten Ergebnissen in diesem Bereich.

4.2.4 Zusammenfassung

Die vorgestellten Untersuchungen zum Gebrauch von Präpositionen und PP in Lernerkorpora (Breindl 2013, Sieradz/Bordag 2014 und Weber 2014) folgen allesamt einem deskriptiven Ansatz, der für die Lernerkorpusforschung typisch ist (Myles 2015: 309, Rankin 2015: 239). Einen direkten Bezug zu aktuellen Annahmen aus der Z/FSE-Forschung, der derzeit in der Lernerkorpusforschung gefordert wird (Granger 2012: 13), lässt sich bei keiner dieser Untersuchungen ermitteln. Die Untersuchungen sind allesamt im Ansatz quantitativ ausgerichtet, die Auswertungen beschränken sich jedoch bestenfalls auf Methoden der deskriptiven Statistik. Insgesamt zeigt sich somit ein Desiderat nach Untersuchungen zum

¹⁵⁹ Für das Schwedische gehen Eeg-Olofsson/Knutsson (2003) entsprechend vor und basieren ihre Entwicklung eines automatischen Fehlererkennungssystems auf die Analyse authentischer Lernerdaten. Ihr regelbasiertes System berücksichtigt auf Basis von 22 Regeln die Fehlertypen *Hinzufügung*, *Auslassung* und *Wahl* im Bereich Präpositionsgebrauch.

¹⁶⁰ Unter Crowdsourcing versteht man eine Art von Arbeitsteilung, bei der (Teil-)Aufgaben (hier: Annotationen von Fehlern) an eine Gruppe Freiwilliger (Nicht-Experten) ausgelagert werden. Sowohl Tetreault/Filatova/Chodorow (2010) als auch Madnani et al. (2011) beziehen sich auf einen Crowdsourcing-Ansatz, bei der Freiwillige gegen einen kleinen Geldbetrag online Aufgaben bearbeiten.

Gebrauch von Präpositionen bzw. PP auf Basis von Lernerkorpusdaten unter Berücksichtigung spracherwerbstheoretischer Erkenntnisse und unter Anwendung statistischer Analyseverfahren. Die vorliegende Dissertation verfolgt das Ziel, zur Schließung dieser Forschungslücke beizutragen.

Die vorgestellten Arbeiten an der Schnittstelle zur maschinellen Sprachverarbeitung machen u.a. deutlich, dass der Entwicklung erfolgreicher Systeme zur automatischen Fehlererkennung im Präpositionsgebrauch sinnvollerweise eine Analyse authentischer, typischer Abweichungen von Lernenden vorausgehen sollte. Anhand von Untersuchungen in Lernerkorpora lassen sich beispielsweise gezielt Schwierigkeiten im Präpositionsgebrauch identifizieren und quantifizieren. Die Ergebnisse können wiederum in der Entwicklung von Systemen zur automatischen Fehlererkennung und -korrektur genutzt werden. Somit bieten linguistische Untersuchungen in Lernerkorpora, so wie sie mit dieser Dissertation vorliegen, wertvolle Anknüpfungspunkte für die Computerlinguistik bzw. für Ansätze zur Entwicklung automatischer Fehlererkennungssysteme für die Zielsprache Deutsch.

5. Zusammenfassung Teil I

Die vorliegende Dissertation widmet sich dem Gebrauch von Präpositionen und Präpositionalphrasen (PP) durch Lernende des Deutschen als Fremdsprache (DaF). Im Vordergrund stehen verschiedene syntaktische Funktionen von PP (Objekt-PP und adverbiale PP) sowie der semantische Gehalt von Präpositionen (semantisierte und desemantisierte Präpositionen). Des Weiteren spielt die Betrachtung von Präpositionen und PP aus kontrastiver Sicht, genauer gesagt die typologische Nähe bzw. Distanz zur Zielsprache Deutsch, eine zentrale Rolle. Meine Terminologie sowie die zentralen Charakteristika von Präpositionen und PP, die für die empirischen Untersuchungen dieser Dissertation von Bedeutung sind, habe ich in Kapitel 1 ausführlich dargelegt. Im Fokus stand dabei u.a. die Verbindung zwischen Form und Funktion von Präpositionen und PP. Die Hauptaussagen hierzu werden an dieser Stelle nochmals wiederholt:

- Adverbiale PP repräsentieren insgesamt zuverlässige Form-Funktionsverbindungen
- Objekt-PP repräsentieren insgesamt weniger zuverlässige Form-Funktionsverbindungen
- Semantisierte Präpositionen (in Objekt-PP) repräsentieren zuverlässige Form-Funktionsverbindungen
- Desemantisierte Präpositionen (in Objekt-PP) repräsentieren weniger zuverlässige Form-Funktionsverbindungen

Der Verbindung zwischen Form und Funktion wird im Rahmen gebrauchsbasierter Ansätze zum Zweit- und Fremdspracherwerb (L2-Erwerb) ein hoher Stellenwert zugeschrieben. In Kapitel 2 habe ich zentrale Annahmen aktueller Theorien zum L2-Erwerb präsentiert, die die empirischen Untersuchungen der vorliegenden Dissertation maßgeblich leiten. Den Annahmen gebrauchsbasierter Ansätze zum L2-Erwerb folgend, habe ich dargelegt, dass Form-Funktionsverbindungen im L2-Erwerb eine zentrale Rolle spielen: Einerseits beeinflusst die Zuverlässigkeit von Form-Funktionsverbindungen in der Zielsprache den Erwerb und

Gebrauch entsprechender sprachlicher Strukturen, andererseits wirken bereits bestehende Form-Funktionsverbindungen (z. B. aus der Erstsprache) auf den Erwerb und Gebrauch von Form-Funktionsverbindungen der Zielsprache ein. Der Gebrauch von Präpositionen und PP durch Deutschlernende wird diesen Annahmen folgend im zweiten Teil der vorliegenden Dissertation empirisch untersucht.

Die Dissertation ist in der Lernerkorpusforschung zu verorten – ein Forschungsfeld, das sich mit der Analyse von Lernerdaten in einem Lernerkorpus beschäftigt. Ein Lernerkorpus wurde in Kapitel 3 definiert als eine systematisch zusammengestellte, digital aufbereitete, mit linguistischen Annotationen angereicherte und elektronisch zugängliche Sammlung von authentischen (schriftlichen) Produktionen von L2-Lernenden. In ihrer interdisziplinären Ausrichtung ist die Lernerkorpusforschung an Schnittstellen zu anderen Forschungsfeldern (u. a. der Zweit- und Fremdspracherwerbsforschung, Z/FSE) zu verorten. Zu den Vorteilen der Lernerkorpusforschung gegenüber traditionellen Datensammlungen und empirischen Zugängen der Zweit- und Fremdspracherwerbsforschung gehört u.a. die Quantität der Daten, die Anwendung korpuslinguistischer Analysemethoden sowie der Anspruch an eine umfassende Dokumentation der Datengrundlage, -erhebung und -auswertung. Die in Lernerkorpora enthaltenen Produktionsdaten eignen sich gut für Untersuchungen zu inter- und intralingualen Einflüssen im Spracherwerb und Sprachgebrauch im Kontext kognitiver Spracherwerbstheorien und bieten somit ein hohes Potenzial für die Z/FSE-Forschung.

In Kapitel 1-3 dieser Dissertation wurden somit das untersuchte Phänomen (Gebrauch von Präpositionen und PP durch Deutschlernende) vorgestellt, die spracherwerbstheoretische Perspektive darauf (gebrauchsbasierte Ansätze) eingenommen sowie der methodologische Zugang bzw. die Verortung im Forschungsfeld (Lernerkorpusforschung an der Schnittstelle zur Z/FSE-Forschung) beschrieben. Die folgende Abbildung (Abb. 15) fasst dies visuell zusammen.

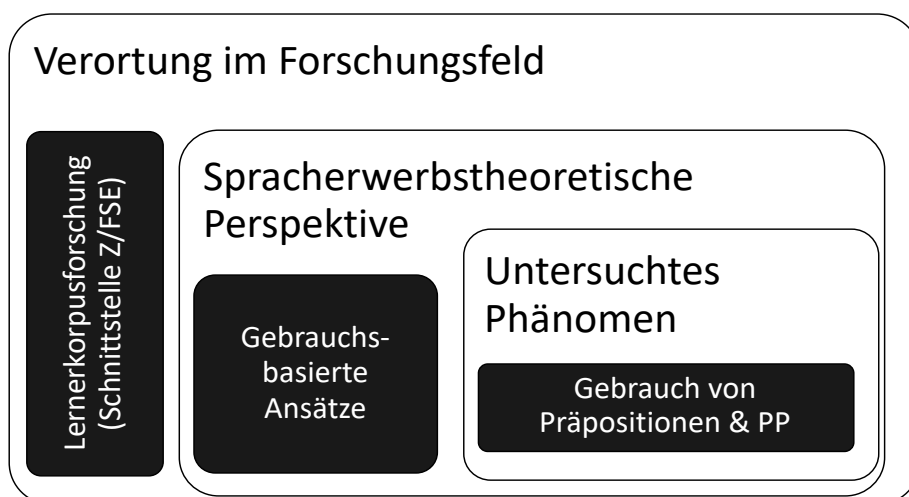


Abbildung 15: Grundlagen der Dissertation: Zusammenfassung.

In Kapitel 4 habe ich den Forschungsstand zum Erwerb und Gebrauch von Präpositionen und PP durch Deutschlernende dargelegt – dabei wurden Erkenntnisse aus der Zweit- und Fremdspracherwerbsforschung (Kap. 4.1) sowie aus der Lernerkorpusforschung berücksichtigt (Kap. 4.2). Im Vordergrund standen dabei (potentielle) Variablen, die den Gebrauch von Präpositionen und PP durch Deutschlernende beeinflussen und welche

Gebrauchsschwierigkeiten diese jeweils zu bedingen scheinen. Im Einzelnen habe ich folgende Einflussvariablen fokussiert und präsentiert: Die Erstsprache der Lernenden, die Sprachkompetenz, die syntaktische Funktion der PP sowie der semantische Gehalt der Präposition. Die zentralen Erkenntnisse aus der Zweit- und Fremdspracherwerbsforschung sind in Kapitel 4.1.6 zusammengefasst. Ich habe gezeigt, dass die Variablen *Erstsprache der Lernenden* und die *Sprachkompetenz* als gut erforschte Variablen gelten können. Für beide Variablen kann ein Einfluss auf den Erwerb und Gebrauch von Präpositionen und PP festgestellt werden. Dahingegen liegen für die Einflussvariablen *Syntaktische Funktion der PP* sowie *Semantischer Gehalt der Präposition* bisher keine aussagekräftigen Befunde vor. In der bisherigen Forschung finden sich jedoch Hinweise auf einen möglichen Einfluss dieser Variablen auf den Gebrauch von Präpositionen und PP.

In Kapitel 4.2 habe ich bisherige Erkenntnisse der Lernerkorpusforschung zum Gebrauch von Präpositionen und PP durch Deutschlernende präsentiert. An der Schnittstelle der Lernerkorpusforschung zur Z/FSE-Forschung liegen bisher wenige bzw. keine empirisch fundierten Erkenntnisse zu möglichen Einflussvariablen im Bereich Präpositionsgebrauch vor. Weiterhin orientieren sich diese wenigen Untersuchungen nicht an aktuellen Annahmen zum L2-Erwerb, was jedoch in der aktuellen Diskussion der Lernerkorpus-Community gefordert wird. Weiterhin liegen zu keiner der in Kapitel 4.1 vorgestellten Einflussvariablen quantitative Untersuchungen vor, die den Einfluss dieser Variablen auf den Gebrauch von Präpositionen und PP durch Deutschlernende sowie dadurch bedingte Gebrauchsschwierigkeiten mittels statistischer Analysen modellieren. Die vorliegende Dissertation trägt daher zur Schließung dieser methodologischen Forschungslücken bei.

An der Schnittstelle der Lernerkorpusforschung zur maschinellen Sprachverarbeitung finden sich Ansätze zur automatischen Erkennung ausgewählter Fehlertypen im Präpositionsgebrauch von Lernenden. Hier liegt der bisherige Schwerpunkt jedoch auf Englisch als Zielsprache. Die Arbeiten sind jedoch insofern für das Deutsche relevant, als sie zeigen, dass Erkenntnisse aus linguistischen Untersuchungen in Lernerkorpora wertvolle Anknüpfungspunkte für die Entwicklung eben solcher Systeme bieten.

Die folgende Abbildung (Abb. 16) fasst die zentralen Aussagen zum in Kapitel 4 dargelegten Forschungsstand zum Gebrauch von Präpositionen und PP durch Deutschlernende visuell zusammen.

Gebrauch von Präpositionen & PP durch Deutschlernende: Forschungsstand

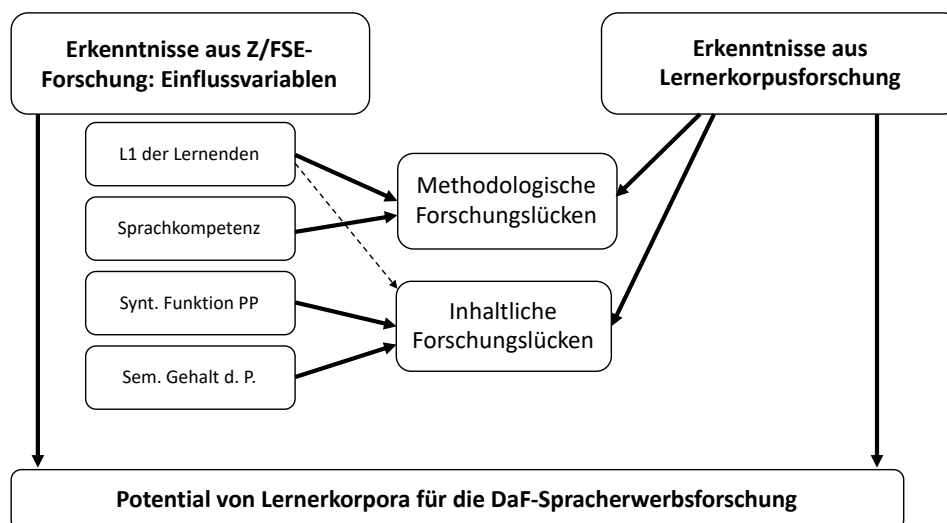


Abbildung 16: Forschungsstand zum Gebrauch von Präpositionen und PP durch Deutschlernende: Zusammenfassung.

Mit meiner Dissertation möchte ich an die bisherigen Forschungserkenntnisse aus der Z/FSE-Forschung sowie der Lernerkorpusforschung anknüpfen und zur Schließung einiger inhaltlicher Forschungslücken beitragen, die sich aus der Darstellung des Forschungsstands ergeben. Ich gehe im Folgenden kurz auf diese inhaltlichen Forschungslücken ein.

In den bisherigen Untersuchungen zum Einfluss der L1 der Lernenden auf den Erwerb und Gebrauch von Präpositionen und PP in der Zielsprache Deutsch dominieren einzelne L1 (vor allem das Türkische) und der Fokus liegt im DaZ-Bereich. Dies bietet mir die Möglichkeit, den Einfluss dieser Variable im Kontext des DaF-Erwerbs zu untersuchen sowie bisher wenig erforschte L1 von Deutschlernenden (z. B. Ungarisch) in den Fokus zu rücken. Hier bietet der empirische Zugang auf Basis von Lernerkorpusdaten eine Möglichkeit, bisherige Erkenntnisse zum Einfluss der L1 auf den Gebrauch von Präpositionen und PP quantitativ zu überprüfen sowie neue Erkenntnisse auf Basis empirischer Untersuchungen zu liefern. Zu den Einflussvariablen *Syntaktische Funktion der PP* sowie *Semantischer Gehalt der Präposition* lassen sich anhand der bisherigen Forschung lediglich Annahmen formulieren. Forschungsarbeiten, die gezielt den Gebrauch von Präpositionen in unterschiedlichen syntaktischen Funktionen von PP untersuchen, liegen meines Wissens nicht vor. Ebenfalls fehlen bisher empirische Untersuchungen zum Einfluss des semantischen Gehalts einer Präposition auf deren Gebrauch bzw. die Ermittlung möglicher Gebrauchsschwierigkeiten bei DaF-Lernenden. Die empirischen Untersuchungen dieser Dissertation, die im folgenden Kapitel präsentiert werden, veranschaulichen insgesamt das Potenzial von Lernerkorpora für die quantitative DaF-Spracherwerbsforschung.

Die obigen Ausführungen motivieren die inhaltliche und methodische Ausrichtung der vorliegenden Dissertation, die in der nachfolgenden Abbildung (Abb. 17) illustriert ist.

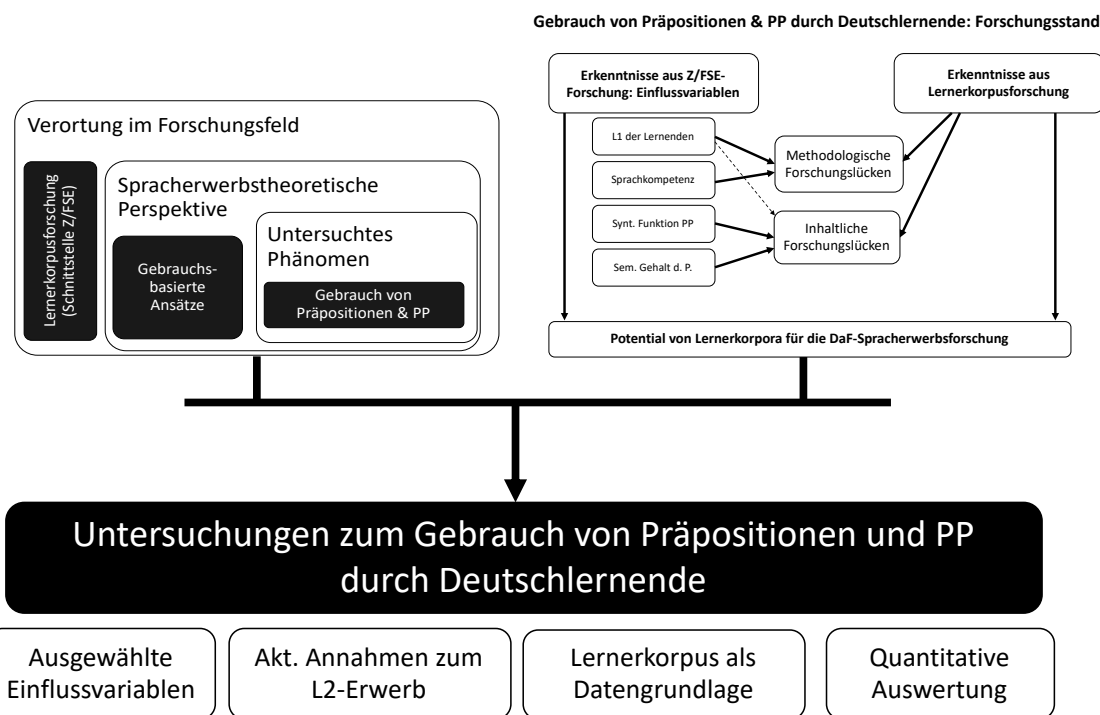


Abbildung 17: Inhaltliche und methodische Ausrichtung der Dissertation.

Im Fokus der Dissertation stehen ausgewählte Einflussvariablen im Gebrauch von Präpositionen und PP, für die ich entsprechende Forschungslücken herausgearbeitet habe. Die Untersuchungen sind erwerbstheoretisch informiert, d. h. sie werden von zentralen aktuellen spracherwerbstheoretischen Annahmen zum L2-Erwerb geleitet. Als Datengrundlage dient ein digitales Lernerkorpus. Der Schwerpunkt liegt auf quantitativen Auswertungen mittels zentraler Analysemethoden der Lernerkorpusforschung¹⁶¹ und statistischen Modellierungen zum Einfluss ausgewählter Variablen auf den Gebrauch von Präpositionen und PP. Die übergeordnete Fragestellung, die im Rahmen einzelner Fallstudien im Lernerkorpus bearbeitet wird, lautet:

Welche Variablen beeinflussen den Gebrauch von Präpositionen und Präpositionalphrasen durch Lernende des Deutschen als Fremdsprache?

Insbesondere liegt der Schwerpunkt auf folgender Forschungsfrage:

Welche der ausgewählten Einflussvariablen bedingen (welche) Gebrauchsschwierigkeiten?

In der vorliegenden Arbeit operationalisiere ich *Gebrauchsschwierigkeit* durch die Korrektheit im Präpositionsgebrauch sowie durch die einzelnen Fehlertypen, die sich im nichtzielsprachlichen Gebrauch von Präpositionen und PP durch DaF-Lernende zeigen.

¹⁶¹ Hierzu zählen die computerunterstützte Fehleranalyse und die kontrastive Interlanguage-Analyse (s. Kap. 3.4).

In den folgenden Kapiteln werden die empirischen Untersuchungen dargelegt, wobei ich zunächst auf die Datengrundlage und die statistischen Auswertungen eingehen werde. Einen Überblick zu den Fallstudien dieser Dissertation gibt Kapitel 7.

Teil II: Empirische Untersuchungen

6. Datengrundlage: Das Lernerkorpus MERLIN

Als Datengrundlage der empirischen Untersuchungen dieser Dissertation dient das Lernerkorpus MERLIN (Abel et al. 2014). Das Lernerkorpus wurde im Rahmen des EU-Projekts „Multilingual Platform for the European Reference Levels: Exploring Interlanguage in Context“ zwischen 2012 und 2014 aufgebaut, wobei unterschiedliche Institutionen (u. a. EURAC Research) am Aufbau des Korpus beteiligt waren¹⁶². Das Lernerkorpus MERLIN enthält schriftliche Texte, die im Rahmen standardisierter Sprachprüfungen durch Sprachlernende unterschiedlicher Kompetenzniveaus produziert wurden. MERLIN ist seit Projektabschluss online über die Plattform <http://www.merlin-platform.eu/> frei zugänglich¹⁶³. Der Aufbau des Lernerkorpus ist ausführlich im Referenzartikel von Abel et al. (2014) beschrieben, weiterhin findet sich eine ausführliche Dokumentation auf der Korpus-Plattform. Im vorliegenden Kapitel möchte ich auf ausgewählte Charakteristiken des Lernerkorpus MERLIN eingehen, die für die vorliegenden Fallstudien relevant sind: Den Korpus-Aufbau und die im Lernerkorpus verfügbaren Daten, die vorliegenden Annotationen im Lernerkorpus sowie auf das Dateiformat und den Korpus-Zugang.

6.1 Korpus-Aufbau und Daten

MERLIN ist ein trilinguales Lernerkorpus und besteht insgesamt aus drei Subkorpora. Entsprechend der Zielsprache (L2) der Sprachlernenden beinhaltet MERLIN ein italienisches, ein tschechisches und ein deutsches Subkorpus. Jedes Subkorpus in MERLIN enthält schriftliche Texte¹⁶⁴, die im Kontext standardisierter Sprachprüfungen des Gemeinsamen Europäischen Referenzrahmen für Sprachen (GeR) entstanden sind und unterschiedliche Kompetenzniveaus von Sprachlernenden repräsentieren. Die Datenerhebung entspricht einem Querschnitt-Design, d. h. „data [are] gathered from different categories of learners from a single point in time“ (Granger 2008: 263). Bei MERLIN handelt es sich gleichzeitig um ein quasi-longitudinales Lernerkorpus (Meunier 2015: 381), da im Korpus unterschiedliche

¹⁶² Die einzelnen Institutionen sind: Die Technische Universität Dresden, die EURAC in Bozen, die Karlsuniversität Prag, das telc-Institut in Frankfurt a.M., das Berufsförderungsinstitut Oberösterreich in Linz sowie die Eberhard-Karls-Universität Tübingen (MERLIN, <http://www.merlin-platform.eu/index.php>) (Mai 2019).

¹⁶³ Ein direkter Zugang zu MERLIN ist auch unter <https://commul.eurac.edu/annis/merlin> möglich (Stand Mai 2019).

¹⁶⁴ Die handschriftlich erhobenen Lernertexte wurden mit dem Editor Xml-mind transkribiert (Näheres s. Abel et al. 2014: 121, zum Xml-Mind-Editor s. <http://www.xmlmind.com>, Mai 2019). Zur Transkription in MERLIN s. MERLIN Richtlinien (2014).

Sprachkompetenzniveaus repräsentiert sind. Im Gegensatz zu einem genuinen longitudinalen Lernerkorpus, das Daten von ein- und denselben Lernenden zu unterschiedlichen Datenerhebungszeitpunkten enthält, beinhaltet ein quasi-longitudinales Lernerkorpus wie MERLIN Daten von unterschiedlichen Lernenden auf unterschiedlichen Kompetenzniveaus (vgl. Gilquin 2015: 14). Abbildung 18 enthält eine Gegenüberstellung des longitudinalen (L) und quasi-longitudinalen (QL) Designs der Datenerhebung, aus der sich ein longitudinales bzw. ein quasi-longitudinales Lernerkorpus ableitet.

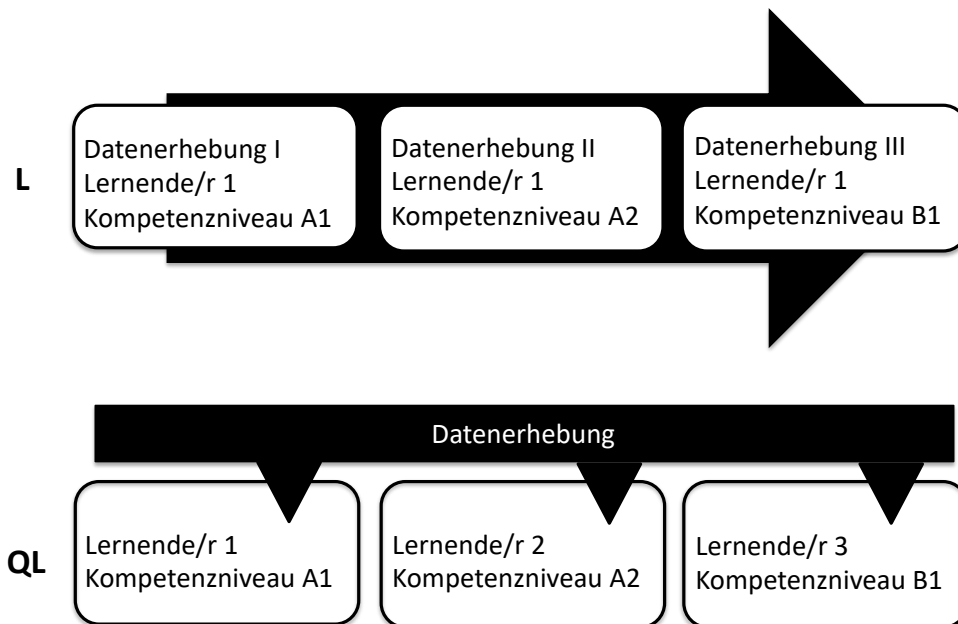


Abbildung 18: Datenerhebungsdesign longitudinal (L) und quasi-longitudinal (QL): Vergleich (nach Meunier 2015: 382).

Ein longitudinales Datenerhebungsdesign (L) bietet die Möglichkeit, die Entwicklung sprachlicher Kompetenzen ausgewählter Individuen präzise nachzuzeichnen und zu interpretieren. Die Erhebung der Daten findet zu unterschiedlichen Zeitpunkten statt (z.T. liegen Jahre dazwischen) und stellt u. a. aus diesem Grund eine Herausforderung dar (vgl. Meunier 2015: 380f., s. auch Kap. 3.2.2.1)¹⁶⁵. Zudem kann die Datenanalyse erst nach Abschluss der gesamten Datenerhebung erfolgen. Die Analyse quasi-longitudinal (QL) gewonnener Lernerdaten lässt lediglich indirekte Aussagen über Entwicklungsverläufe und –muster einer Gruppe von Lernenden zu. Dennoch ist es möglich, durch Analysen in einem quasi-longitudinalen Lernerkorpus entwicklungsbedingte Aussagen über die gesamte Gruppe von Lernenden zu treffen (Meunier 2015: 382). Meunier (2015) unterstreicht den Mehrwert von Lernerkorpora, die unterschiedliche Kompetenzniveaus repräsentieren, für die Erforschung von Erwerbsverläufen. Nach Meunier (2015: 396) sind „[l]earner corpora [...] solid and reliable data sources to trace learners’ proficiency development in an L2.“ Bei MERLIN handelt es sich um ein quasi-longitudinales Lernerkorpus mit zuverlässig ermittelten Kompetenzniveaus, das sich für die Untersuchung von Erwerbsverläufen optimal eignet.

¹⁶⁵ Eine weitere Herausforderung besteht darin, dass nicht alle Lernenden, die zu Beginn für die Untersuchung ausgewählt wurden, zu allen Erhebungszeitpunkten zur Verfügung stehen, sodass bei einem longitudinalen Design i.d.R. mit ‚Datenschwund‘ zu rechnen ist (Meunier 2015: 381).

Im Folgenden gehe ich lediglich auf das deutsche Subkorpus von MERLIN ein und verweise fortan mit *MERLIN* lediglich auf dieses deutsche Subkorpus. Das Lernerkorpus MERLIN lässt sich in Bezug auf die Lernerkorpus-Typologie nach Granger (2008) gemäß Abbildung 19 einordnen (Näheres zur Typologie s. Kap. 3.2.2).

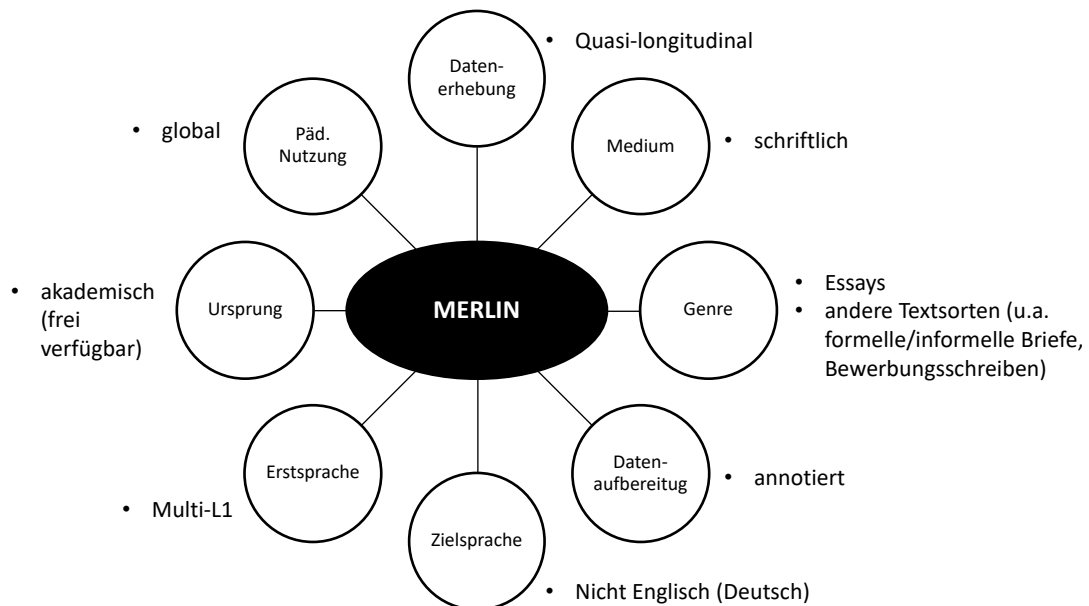


Abbildung 19: Einordnung des Lernerkorpus MERLIN in die Lernerkorpus-Typologie.

MERLIN verfolgt das primäre Ziel, die Kompetenzniveaus des GeR anhand authentischer Sprachdaten von Sprachlernenden zu illustrieren und somit zur Erforschung und Validierung der GeR-Kompetenzniveaus beizutragen (Abel et al. 2014: 111f.)¹⁶⁶. MERLIN enthält authentische¹⁶⁷ Texte von DaF-Lernenden (MERLIN-Projekt 2014: 3), die im Rahmen von standardisierten Sprachtests unterschiedlicher Kompetenzniveaus produziert wurden. Die Sprachtests haben einen Bezug zum GeR (Europarat 2001) und wurden an renommierten Institutionen entwickelt. Sie entsprechen allesamt den Standards der *Association of Language Testing in Europe (ALTE)*¹⁶⁸ (Abel et al. 2014: 113). Die Kompetenzniveaus des GeR lassen sich in unterschiedliche qualitative Klassen einteilen (vgl. Abb. 20). Die Kompetenzniveaus A1, A2, B1, B2, C1 und C2 werden auch *Kriterien-Niveaus* genannt, während die Kompetenzniveaus A2+, B1+ und B2+ *Plus-Niveaus* genannt werden und zwischen den Kriterien-Niveaus A2, B1 und B2 zu verorten sind (vgl. Abb. 20) (Europarat 2001: 41). Obwohl Abbildung 20 suggeriert, dass die Plus-Niveaus keine direkte Verbindung zu den entsprechenden Kriterien-Niveaus

¹⁶⁶ Da der GeR sprachübergreifend angelegt ist, sind die Beschreibungen einzelner Kompetenzniveaus entsprechend allgemein verfasst. Die fehlende Sprachspezifik des GeR, aber auch seine fehlende empirische Grundlage wurden vielerorts kritisiert (vgl. z. B. Abel 2010: 206ff., Little 2006: 178ff.). Seit längerer Zeit wird seitens des Europarats die Ausarbeitung komplementierender Instrumente zur Illustration einzelsprachspezifischer Merkmale der GeR-Kompetenzniveaus gefördert (Abel et al. 2014: 112), siehe z. B. Glaboniat (2010) für das Deutsche.

¹⁶⁷ ‚Authentisch‘ in dem Sinne, als dass es sich um Lernertexte handelt, die nicht unter Experimentbedingungen entstanden sind.

¹⁶⁸ <http://www.alte.org/> (Mai 2019).

aufweisen, sind die Plus-Niveaus A2+, B1+ und B2+ als den entsprechenden Kriterien-Niveaus zugehörig zu verstehen, d. h. das Plus-Niveau A2+ ist dem Kriterien-Niveau A2 zugeordnet (vgl. Europarat 2001: 41).

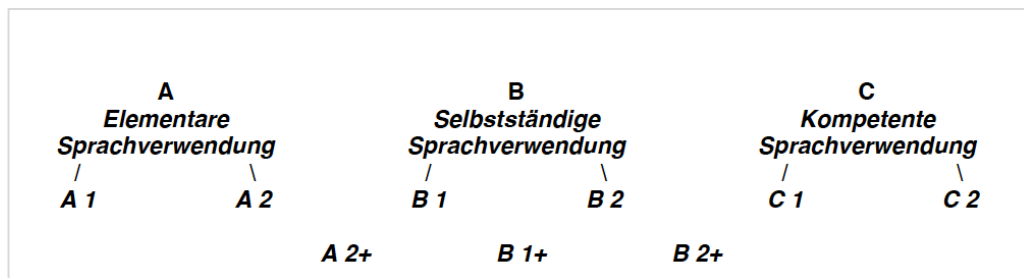


Abbildung 20: Kompetenzniveaus gemäß dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GeR) (Europarat 2001: 41).

Die schriftlichen Texte in MERLIN wurden im Rahmen von Sprachtests für die Kriterien-Niveaus A1-C1 erhoben (=GeR-Testniveau). Je GeR-Testniveau liegen drei verschiedene Testaufgaben vor. Die Aufgabenstellungen enthielten alltagsbezogene Themen, zu denen die Lernenden einen schriftlichen Text produzieren sollten. Die Texte repräsentieren unterschiedliche Textsorten (Brief an Freund/-in, Bewerbungsschreiben, Bericht, Essay) und unterscheiden sich damit auch in ihrer Formalität. Des Weiteren sind in Abhängigkeit vom GeR-Testniveau die Textproduktionen in unterschiedlichem Maße angeleitet. Zur Veranschaulichung dienen die folgenden Abbildungen mit exemplarischen Beispielen für Aufgabenstellungen der GeR-Testniveaus A2 (Abb. 21) und C1 (Abb. 22) (s. MERLIN¹⁶⁹).

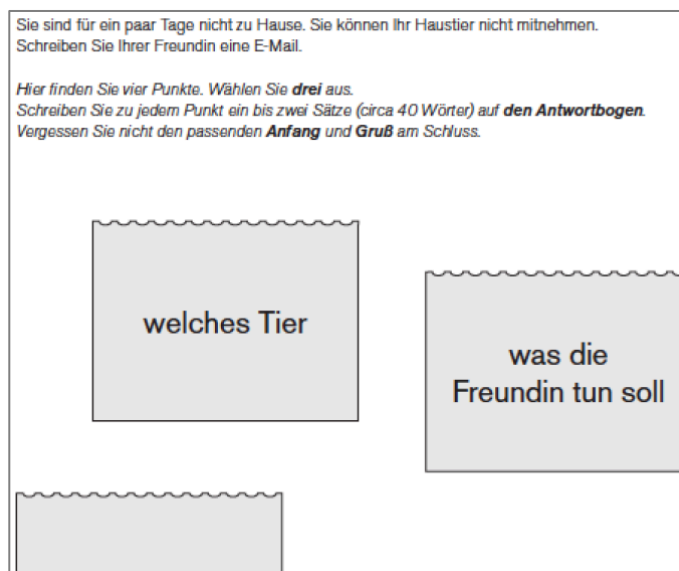


Abbildung 21: Aufgabenstellung GeR-Testniveau A2: Beispiel (Ausschnitt).

¹⁶⁹ http://www.merlin-platform.eu/C_mcorpus.php#anchor3 (Mai 2019).

In einem Internetforum ging es um das Thema: „Ist es sinnvoll, auch im Ausland die Traditionen seines Heimatlandes fortzusetzen, oder sollte man sich an der Kultur des Gastlandes orientieren?“ Sie möchten Ihre Meinung dazu erläutern.
Schreiben Sie einen Artikel zu diesem Thema (ca. 200 Wörter)!

Abbildung 22: Aufgabenstellung GeR-Testniveau C1: Beispiel (Ausschnitt).

Die Gemeinsamkeit der Aufgabenstellungen in MERLIN besteht darin, dass sie, anders als es z. B. bei Übersetzungsaufgaben oder Lückentexten der Fall wäre, den Lernenden (in unterschiedlichem Maße) ermöglichen, ihre Textproduktionen sprachlich relativ frei zu gestalten (vgl. Gilquin 2015: 10). Somit liegen in MERLIN lernersprachliche Produktionen vor, die nach Granger (2008: 261) in einem Kontinuum der ‚Natürlichkeit‘ recht weit oben angesiedelt werden können und sich daher gut für Untersuchungen im Kontext der Z/FSE-Forschung eignen.

Je GeR-Testniveau sind in MERLIN ca. 200 schriftliche Texte enthalten, wobei die Tokenanzahl je GeR-Testniveau variiert. Der Gesamtumfang liegt bei 1033 Texten bzw. ca. 155000 Token (s. Tab. 18)¹⁷⁰. Damit enthält MERLIN schriftliche Produktionen von insgesamt 1033 DaF-Lernenden.

GeR-Testniveau	Anzahl Lernertexte	Anzahl Token
A1	206	13142
A2	209	18332
B1	210	27820
B2	204	40949
C1	204	54092
Gesamt	1033	154335

Tabelle 18: Verteilung der Lernertexte und Token in MERLIN je GeR-Testniveau: Übersicht.

Die im Rahmen der standardisierten Sprachtests produzierten Lernertexte wurden nach der Erhebung individuell hinsichtlich der Kompetenzniveaus des GeR durch ausgebildete RaterInnen bewertet (=GeR-Gesamtniveau) (vgl. Abel et al. 2014: 113). Somit liegen in MERLIN pro Lernertext zwei GeR-bezogene Bewertungen bzw. Kompetenzniveaus vor: Das *GeR-Testniveau* und das *GeR-Gesamtniveau*. Dabei bezieht sich das GeR-Testniveau auf das GeR-Niveau des absolvierten Sprachtests während das GeR-Gesamtniveau auf die im Anschluss vorgenommene Bewertung des produzierten Textes referiert. Ein Lernertext kann damit dem GeR-Testniveau A2 zugeordnet sein, der produzierte Lernertext jedoch z. B. entsprechend eines niedrigen (A1), höheren (B1) oder demselben GeR-Gesamtniveaus (A2) bewertet sein. Die folgende Tabelle (Tab. 19) veranschaulicht diesen Unterschied zwischen GeR-Testniveau und GeR-Gesamtniveau.

¹⁷⁰ Damit ist das Lernerkorpus MERLIN ein vergleichsweise kleines Lernerkorpus. Zum Vergleich: Das International Corpus of Learner English (ICLE) (v2) enthält insgesamt 3,7 Mio. Token (Granger et al. 2009).

Benennung	Grundlage der Zuordnung zum GeR-Niveau	Erläuterung
GeR-Testniveau	Merkmale des zugrundeliegenden Sprachtests	Person X absolviert erfolgreich einen Sprachtest des Niveaus A2. Das Kompetenzniveau von Person X entspricht A2.
GeR-Gesamtniveau	Merkmale des individuell produzierten Textes	Person X absolviert erfolgreich einen Sprachtest des Niveaus A2. Der produzierte Text wird separat bewertet und erhält das Niveau B1. Das Kompetenzniveau von Person X entspricht B1.

Tabelle 19: GeR-Testniveau und GeR-Gesamtniveau: Gegenüberstellung.

Die (Nach-)Bewertung der produzierten Lernertexte bedingt im Lernerkorpus eine Verteilung der GeR-Gesamtniveaus, die von der Verteilung der GeR-Testniveaus (vgl. Tab. 18) abweicht. Die Verteilung der GeR-Gesamtniveaus in MERLIN ist in Tabelle 20 aufgeführt. Da die geschulten RaterInnen bei der (Nach-)Bewertung der Texte auf alle GeR-Kompetenzniveaus zurückgreifen konnten, sind in der neuen Verteilung neben den Kriterien-Niveaus A2, B1, etc. ebenfalls die Plus-Niveaus A2+, B1+ etc. vertreten.

GeR-Gesamtniveau			Anzahl Lernertexte	Anzahl Token	
A1			57	2505	
A2	A2	199	306	13353	23538
	A2+	107		10195	
B1	B1	217	332	25717	47504
	B1+	115		21787	
B2	B2	219	292	50477	68386
	B2+	73		17909	
C1			42	11251	
C2			4	1151	
Gesamt			1033	154335	

Tabelle 20: Verteilung der Lernertexte und Token in MERLIN je GeR-Gesamtniveau: Überblick.¹⁷¹

Die folgende Abbildung (Abb. 23) veranschaulicht die Gegenüberstellung der (prozentualen) Verteilung von Lernertexten je GeR-Testniveau und GeR-Gesamtniveau in MERLIN. Wie aus der Abbildung deutlich wird, liegen im Korpus vorwiegend Lernertexte vor, die den GeR-Niveaus A2, B1 und B2 zugeordnet wurden. Damit ergibt sich eine ungleiche Verteilung der Daten auf die einzelnen Kompetenzniveaus, die sich ebenfalls auf die Datengrundlage der einzelnen Fallstudien dieser Dissertation auswirkt.

¹⁷¹ Anmerkung zur Tabelle: Die Aufteilung unterscheidet sich von der Aufteilung in Abel et al. (2014), sie entspricht aber der Aufteilung der Texte, so wie sie in MERLIN (Stand März 2017) ermittelt wurden.

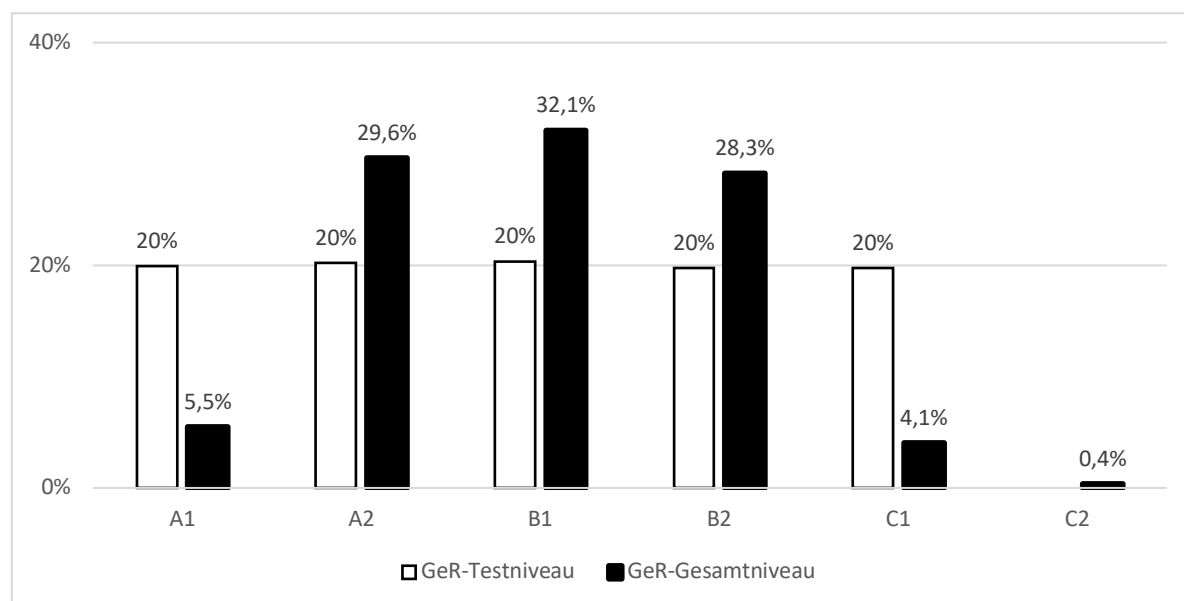


Abbildung 23: Verteilung der Lernertexte in MERLIN je GeR-Testniveau und GeR-Gesamtniveau: Vergleich (prozentual).

Die Zuordnung von Kompetenzniveaus auf Basis der Bewertung individueller Lernertexte entspricht textzentrierten Methoden zur Ermittlung von Kompetenzniveaus in Lernerkorpora (vgl. Carlsen 2012: 165f.). Textzentrierte Methoden fußen auf internen Kriterien zur Bewertung des Kompetenzniveaus, wie z. B. dem Ergebnis in einem Einstufungstest oder die Bewertung durch geschulte RaterInnen (vgl. Carlsen 2012: 166.), wie dies z. B. bei MERLIN vorliegt. Lernerzentrierte Methoden dahingegen ermitteln das Kompetenzniveau von Lernenden anhand externer Kriterien, z. B. auf Basis der Sprachlernjahre oder auf Basis von allgemeinen, teils subjektiven Einschätzungen durch Sprachlehrende ohne Rückgriff auf standardisierte Verfahren. Lernerzentrierte Methoden zur Bestimmung der Sprachkompetenz bzw. des Kompetenzniveaus sind in der Lernerkorpusforschung derzeit zwar verbreitet, werden jedoch aufgrund ihrer begrenzten Aussagekraft auch kritisch hinterfragt (vgl. Callies 2015: 52), da sie u. U. eine Varianz der sprachlichen Leistungen von Lernenden innerhalb eines extern ermittelten Kompetenzniveaus verdecken, was sich wiederum negativ auf linguistische Untersuchungen im Korpus auswirken kann (vgl. Carlsen 2012: 164). Durch die Anwendung textzentrierter Methoden zur Ermittlung des Kompetenzniveaus von Lernenden bzw. Lernertexten bietet das Lernerkorpus MERLIN somit ein valides Metadatum zur sprachlichen Kompetenz von Lernenden. Aus diesem Grund eignet sich MERLIN sehr gut für die Untersuchung von Sprachentwicklungsverläufen. In den empirischen Untersuchungen der vorliegenden Dissertation wird ausschließlich das GeR-Gesamtniveau berücksichtigt. Das bedeutet jedoch gleichzeitig, dass die GeR-Gesamtniveaus A2, B1 und B2 im Vergleich zu den anderen GeR-Gesamtniveaus in der Datengrundlage der einzelnen Fallstudien überrepräsentiert sind (vgl. Abb. 23).

Neben dem GeR-Test- und -Gesamtniveau sind in MERLIN weitere Metadaten zu den Lernertexten bzw. Lernenden verfügbar. Eine Übersicht über die verfügbaren Metadaten in MERLIN gibt die folgende Tabelle (Tab. 21). Für die empirischen Untersuchungen der

vorliegenden Dissertation sind vor allen die Metadaten *Test-Aufgabe*, *Erstsprache der Lernenden*¹⁷² sowie das *GeR-Gesamtniveau* relevant.

Metadatum		Erläuterung
Lernender/r	Erstsprache (L1)	Die Erstsprache des/der Lernenden (z. B. Türkisch)
	Alter	Das Alter des/der Lernenden, in Altersbereichen: 12-19 Jahre, 20-29 Jahre, 30-59 Jahre, 60-75 Jahre o. keine Angabe
	Geschlecht	Das Geschlecht des/der Lernenden (w/m)
Lernertext	Text-ID-Nummer	Eine Identifikationsnummer des Lernertexts bzw. des/der Textproduzenten/-in (z. B. 1091_0000155) ¹⁷³ .
	Thema Test-Aufgabe	Thema der zugrundeliegenden Textproduktionsaufgabe (z. B. <i>Freundin um Tierhüten bitten</i>)
	GeR-Testniveau	Das GeR-Niveau des zugrundeliegenden Sprachtests (A1, A2, B1, B2, C1)
	GeR-Gesamtniveau ¹⁷⁴	Das nachträglich auf Basis des produzierten Textes ermittelte GeR-Niveau (A1, A2, A2+, B1, B1+, B2, B2+, C1, C2)

Tabelle 21: Metadaten in MERLIN: Überblick.

Die vollständigen Aufgabenstellungen, die den Lernertexten zugrunde liegen, stehen online zur Verfügung. Je GeR-Testniveau handelt es sich dabei um drei verschiedene Aufgabenstellungen. Die Rahmenbedingungen (z. B. Zeitbeschränkung) sind jeweils bei den einzelnen Aufgaben dokumentiert (MERLIN¹⁷⁵). Durch die Auswahl bestimmter Metadaten bzw. Kombinationen von Metadaten können in MERLIN spezifische Subkorpora erstellt werden. Die Metadaten können ebenfalls für die Datenauswertungen berücksichtigt werden.

6.2 Annotationen

In diesem Abschnitt gehe ich auf die für die empirischen Untersuchungen relevanten Annotationsebenen in MERLIN ein.

6.2.1 Zentrale Annotationsebenen: Überblick

MERLIN verfügt über eine Mehr-Ebenen-Architektur (Lüdeling et al. 2005, Reznicek/Lüdeling/Hirschmann 2013), in der automatische und manuelle Annotationen der Lernertexte integriert sind. Tabelle 22 gibt einen Überblick über die zentralen automatischen

¹⁷² Die Erstsprache ist für den Großteil der Lernenden dokumentiert.

¹⁷³ Die ID wurde jedem Lernertext, jeder/m Lernenden im Rahmen des Anonymisierungsprozesses bei der Digitalisierung bzw. Transkription der handschriftlichen Texte zugewiesen, s. MERLIN Richtlinien (2014: 3).

¹⁷⁴ Je Bewertungskriterium liegen hier ebenfalls Einzel-Bewertungen vor, z. B. zum Wortschatzspektrum oder zur Kohärenz (s. https://merlin-platform.eu/C_mcorpus.php#anchor5, Mai 2019).

¹⁷⁵ https://merlin-platform.eu/C_mcorpus.php#anchor3 (Mai 2019).

Annotationen und Annotationsebenen. Zu den eingesetzten Annotationswerkzeugen s. MERLIN¹⁷⁶).

Annotationsebene	Erläuterungen
Tokenisierung	Zerlegung eines Texts in einzelne Einheiten (=Token) (Token können z. B. Wortformen oder Satzzeichen sein)
Lemmatisierung	Annotation der Grundform einzelner Token (Wortformen)
Wortarten (Part of Speech, kurz POS)	Annotation der Wortarten (gemäß dem Stuttgart-Tübingen-Tagset (STTS), s. Schiller/Teufel/Stöckert 1999)
Dependenzstrukturen (Parsing)	Annotation der vorliegenden syntaktischen Strukturen

Tabelle 22: Zentrale automatische Annotationen in MERLIN: Überblick.

Wie in Tabelle 22 aufgeführt ist, gehören zu den automatischen Annotationen u.a. die Tokenisierung, die Lemmatisierung, die Part-of-Speech-Annotationen (POS-Annotationen) sowie die Annotation von Dependenzstrukturen. Da die verwendeten automatischen Annotationswerkzeuge auf standardsprachlichen Sprachdaten trainiert sind, operieren diese weniger präzise auf Lernertexten (s. Kap 3.3.1). Zusätzlich muss beachtet werden, dass Lernertexte neben nicht-zielsprachlichen orthographischen und grammatischen Formen ebenfalls Auslassungen enthalten können, d. h. fehlende Artikel oder Präpositionen sind in MERLIN nicht durch automatische Annotationswerkzeuge erfasst. Siehe das folgende Beispiel (Abb. 24) zur Illustration.

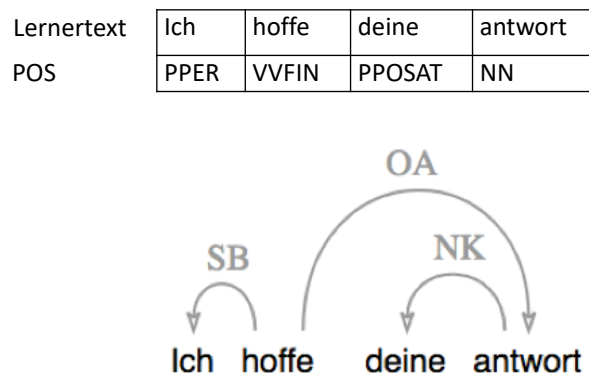


Abbildung 24: Beispiel für automatische Annotationen in MERLIN: Part-of-Speech (POS) (oben) und Dependenzstrukturen (unten)¹⁷⁷.

Wie man Abbildung 24 entnehmen kann, wurde die im Lernertext ausgelassene Präposition *auf* bei der automatischen Annotation nicht erfasst – dies spiegelt sich in der fehlenden POS-Annotation (Ebene POS) dieser Einheit wider sowie in der Annotation der syntaktischen Strukturen. Durch die ausgelassene Präposition wird die vom Lernenden realisierte NP *deine Antwort* als Akkusativobjekt (OA) interpretiert, obwohl das hier vom Lernende verwendete

¹⁷⁶ https://merlin-platform.eu/C_research.php#anchor233 (Mai 2019).

¹⁷⁷ PPER = Personalpronomen, VVFIN = finites Verb, PPOSAT = Possessivpronomen, NN = normales Nomen. SB=Subjekt, OA=Akkusativobjekt, NK=Kern der Nominalphrase. ID: 1091_0000275.

Verb *hoffen* eine Objekt-PP regiert. Bei der automatischen Annotation nicht erfasste Einheiten können in MERLIN somit nicht direkt mit Rückgriff auf entsprechende Annotationsebenen ermittelt werden.

Neben automatischen Annotationen liegen in MERLIN ebenfalls manuelle Annotationen zu unterschiedlichen linguistischen Bereichen vor. Die folgende Tabelle (Tab. 23) gibt einen Überblick über die für die empirischen Untersuchungen relevanten manuellen Annotationen bzw. Annotationsebenen in MERLIN (s. dazu auch MERLIN project 2014). Die im Korpus vorliegenden (engl.) Bezeichnungen der einzelnen Annotationsebenen sind jeweils in Klammern angegeben. Für die einzelnen Annotationswerkzeuge siehe Abel et al. (2014: 121) und MERLIN¹⁷⁸.

Annotationsebene	Erläuterungen
ZH1 (TH1) (minimale Zielhypothese)	Zielsprachliche Rekonstruktion von orthographischen und grammatischen Abweichungen in der Original-Lerneräußerung.
ZH1Diff (TH1Diff) (tlw. automatische Erzeugung; ggf. manuelle Korrektur)	Annotationen der Differenzen zwischen Original-Lerneräußerung und ZH1. Die zentralen Editier-Tags lauten: Einfügung (INS), Änderung (CHA), Löschung (DEL), Verschiebung (MOVS/T).
Fehlerannotationen (EA1)	Annotationen der zielsprachlichen Abweichungen zw. Original-Lerneräußerung und ZH1 im Bereich Grammatik und Orthographie mit Rückgriff auf vordefinierte Annotationskategorien (Wisniewski et al. 2014).

Tabelle 23: Zentrale manuelle Annotationen in MERLIN: Überblick.

Im Folgenden gehe ich näher auf die in Tabelle 23 angegebenen Annotationsebenen ein.

6.2.2 Zielhypothesen

Eine Zielhypothese (ZH) ist eine zielsprachliche Rekonstruktion einer abweichenden Äußerung (Lüdeling 2008: 126). Eine ZH dient als Normalisierungsebene im Korpus (vgl. Lemnitzer/Zinsmeister 2015: 85f.) und macht die Auszeichnung von Fehlern im Korpus explizit und nachvollziehbar (vgl. Lüdeling 2008: 135, Näheres s. auch Kap. 3.3.2). Bei den Richtlinien zur Erstellung der ZH lehnt sich MERLIN an das Lernerkorpus Falko an (Lüdeling et al. 2008, Reznicek et al. 2012). In MERLIN liegen, äquivalent zu Falko, zwei ZH vor, die je verschiedene linguistische Annotationsebenen repräsentieren. Die minimale Zielhypothese (ZH1)¹⁷⁹ beinhaltet Rekonstruktionen auf der Ebene der Orthographie und Grammatik, während in der erweiterten Zielhypothese (ZH2) zusätzlich Abweichungen auf der lexikalischen, textstrukturellen und pragmatischen Ebene rekonstruiert sind. Für die empirischen Untersuchungen dieser Dissertation ist lediglich die Ebene der ZH1 von Relevanz. Daher gehe

¹⁷⁸ https://merlin-platform.eu/C_research.php#anchor231 (Mai 2019).

¹⁷⁹ Auch TH genannt, für engl. *target hypothesis*.

ich im Folgenden nicht weiter auf die Ebene ZH2 ein¹⁸⁰. Abbildung 25 illustriert die Ebene der ZH1 und ZH1Diff in MERLIN anhand eines Beispiels. Wie man Abbildung 25 entnehmen kann, kodiert die Ebene ZH1Diff jeweils die Differenzen zwischen der Original-Lerneräußerung und der ZH1 (siehe auch MERLIN project 2014: 15, 45). Die Differenz ist dabei mittels Editiertags (hier: INS) repräsentiert. Das Kürzel INS kodiert die Einfügung eines Elements (INS, engl. *insert*), hier die Präposition *auf* (s. Abb. 25).

O	Ich	warte		Ihre	Antwort	mit	Ungeduldigkeit
ZH1	Ich	warte	auf	Ihre	Antwort	mit	Ungeduldigkeit
ZH1Diff			INS				

Abbildung 25: Die Ebene ZH1Diff in MERLIN: Beispiel 1. O=Original-Lerneräußerung.

Ein weiteres zentrales Editiertag in MERLIN ist DEL (engl. *delete*) zur Kodierung gelöschter Token, vgl. dazu die folgende Abbildung (Abb. 26). Näheres zu Editiertags s. Kapitel 3.3.2.

O	Ich	möchte	auf	eine	baldige	Antwort
ZH1	Ich	möchte		eine	baldige	Antwort
ZH1Diff			DEL			

Abbildung 26: Die Ebene ZH1Diff in MERLIN: Beispiel 2. O=Original-Lerneräußerung.

In der Phase der Datenaufbereitung wurde die ZH1Diff-Ebene mithilfe des Falko-Excel-Add-Ins (Reznicek 2012) automatisch erstellt und ggf. anschließend manuell korrigiert.

Die ZH1 bildet die Basis für die Fehlerannotationen im Bereich der Grammatik und Orthographie. Wie oben erwähnt, handelt es sich bei der ZH um eine zielsprachliche Rekonstruktion einer abweichenden Äußerung. Da abweichende Lerneräußerungen i.d.R. unterschiedliche Möglichkeiten der zielsprachlichen Rekonstruktion zulassen, ist die Angabe einer expliziten ZH im Lernerkorpus umso wichtiger. Sie gewährleistet sowohl eine transparente und nachvollziehbare Fehlerannotation als auch eine systematische und konsistente Fehlerannotation. Des Weiteren ist sie eine Voraussetzung für die Replizierbarkeit der Datenauswertungen in einem Lernerkorpus (s. auch die Ausführungen in Kap. 3.3.2). Für die Erstellung der ZH1 in MERLIN liegen detaillierte Richtlinien vor (MERLIN project 2014: 14ff.).

6.2.3 Fehlerannotationen

Die in MERLIN vorliegenden Fehlerannotationen spiegeln die Abweichung zwischen der Original-Lerneräußerung und der ZH1 wider. Die Definition von *Fehler*, die in der Phase der Datenaufbereitung in MERLIN konstitutiv war, ist die Definition eines Fehlers nach Lennon (1991) (MERLIN project 2014: 15¹⁸¹, Wisniewski et al. 2014: 27). Lennon (1991: 182) definiert einen Fehler als

¹⁸⁰ Die Annotationsebene ZH2 liegt lediglich für das Kernkorpus in MERLIN vor, das die GeR-Gesamtniveaus A2 und B2 umfasst (vgl. MERLIN project 2014: 3). Die Ebenen ZH1 und ZH1Diff liegen für jeden Lernertext in MERLIN vor.

¹⁸¹ Dieselbe Fehlerdefinition findet man auch im Lernerkorpus *Falko* (s. Reznicek et al. 2012: 38).

a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers' native speaker counterparts.

Die ZH1 in MERLIN kann Lenbons (1991: 182) Definition folgend als „native speakers' counterpart“ angesehen werden. Ein Fehler liegt demnach immer dann vor, wenn die zielsprachliche Rekonstruktion (ZH1-Ebene) Abweichungen von der Ebene der Original-Lerneräußerung aufweist. Ein Fehler ist somit in MERLIN, und für die Fallstudien der vorliegenden Dissertation, formal definiert als eine Abweichung zwischen der Original-Lerneräußerung und der entsprechenden ZH1.

Die Fehlerannotationen in MERLIN wurden manuell erstellt und mithilfe der Annotationssoftware MMAX2 (Müller/Strube 2006) in die digitalen Lernertexte eingefügt. Da für Lernerkorpora derzeit (noch) kein Fehlerannotationsstandard existiert, orientiert sich das Fehlerannotationssystem eines Lernerkorpus daher im Allgemeinen vorwiegend am Forschungsinteresse, das mit dem Korpus verfolgt wird (vgl. Díaz-Negrillo/Domínguez-Fernández 2006). Unter einem Fehlerannotationssystem verstehe ich an dieser Stelle die Fehlerannotationskategorien mit den entsprechenden Fehlertags und Hinweise zur Vergabe dieser Tags. Da MERLIN das Ziel verfolgt, die Kompetenzniveaus des GeR anhand authentischer Daten zu illustrieren und somit zur Erforschung und Validierung der Kompetenzniveaus beizutragen, ist das Annotationssystem in MERLIN entsprechend umfangreich und umfasst unterschiedliche linguistische Bereiche (vgl. Wisniewski et al. 2014). Insgesamt enthält das Fehlerannotationssystem von MERLIN 65 Annotationskategorien, d. h. 65 Tags, die u.a. aus den kompetenzbezogenen Indikatoren des GeR sowie auf Basis extensiver Literaturrecherchen zu linguistischen Untersuchungen im Bereich der Spracherwerbs- und der Sprachtestforschung abgeleitet wurden. Zudem fanden die Ergebnisse von Lehrbuchanalysen sowie induktiv hergeleitete linguistische Merkmale Berücksichtigung im Fehlerannotationssystem von MERLIN (vgl. Abel et al. 2014: 115ff.).

Das Annotationssystem von MERLIN umfasst Fehlertags in den Bereichen Grammatik und Orthographie (Bezug zu Ebene ZH1) sowie in den Bereichen Lexik, Kohärenz und Kohäsion, sprachlicher Stil und Pragmatik (Bezug zu Ebene ZH2) (vgl. MERLIN project 2014, Wisniewski et al. 2014). Entsprechend des Bezugs auf die beiden Ebenen ZH1 und ZH2 ergeben sich auch zwei Ebenen der Fehlerannotation (engl. Error Annotation, kurz EA): *EA1* und *EA2*. Dabei enthält die Ebene *EA1* ausschließlich Fehlerannotationen im Bereich der Grammatik (hierzu zählt auch der Gebrauch von Präpositionen) und Orthographie. Ich möchte im Folgenden lediglich auf diese Ebene eingehen, da diese für die empirischen Untersuchungen im Lernerkorpus relevant ist. Insbesondere möchte ich die Fehlerannotationskategorien im Bereich Präpositionsgebrauch fokussieren.¹⁸²

¹⁸² MERLIN erkennt die Lernaltersprache als eigenständiges Sprachsystem an (Selinker 1972) und kodiert neben den sprachlichen Abweichungen ebenfalls die Existenz anderer, fehlerunabhängiger Merkmale der Lernaltersprache. So findet man neben den Fehlerannotationen in MERLIN auch die Auszeichnung lernaltersprachlicher Merkmale, die unabhängig von ihrer Korrektheit erfasst bzw. annotiert werden. Es handelt sich dabei um so genannte *existence tags*, die z. B. die Realisierung von Sprechakten oder den

Die Ebene EA1 liegt für den Großteil – ca. 73% – der Texte im Korpus vor. Folglich enthalten ca. 24% aller Lernertexte keine Fehlerannotationen und damit auch keine Fehlerannotationen im Bereich Präpositionsgebrauch (MERLIN¹⁸³). Für die empirischen Untersuchungen in dieser Dissertation war somit teilweise eine manuelle, nachträglich vorgenommene Fehlerannotationen im Bereich Präpositionsgebrauch notwendig¹⁸⁴. Die nachfolgende Tabelle (Tab. 24) gibt einen Überblick über die Fehlerannotationskategorien und die entsprechenden Fehlertags im Bereich Präpositionsgebrauch im MERLIN-Korpus, die für die empirischen Untersuchungen zentral sind. In den empirischen Untersuchungen entsprechen die Fehlertypen im Präpositionsgebrauch den hier präsentierten Fehlerannotationskategorien. Eine detaillierte Darstellung aller Fehlerannotationskategorien in MERLIN bietet das Fehlerannotationsschema von MERLIN (Wisniewski et al. 2014).

Linguistischer Bereich: Grammatik	Fehlerannotationskategorien	Fehlertag
Präpositionen	1. Auslassung einer Präposition 2. Hinzufügung einer Präposition 3. Wahl der Präposition 4. Position der Präposition 5. Zusammenziehung der Präposition 6. Trennung der Präposition	1. G_Prep_O 2. G_Prep_Ad 3. G_Prep_Ch 4. G_Prep_Pos 5. G_Prep_Merge 6. G_Prep_Split

Tabelle 24: Fehlerannotationskategorien und Fehlertags im Bereich Präpositionsgebrauch in MERLIN (Wisniewski et al. 2014: 12).

Die Fehlertags der einzelnen Fehlerannotationskategorien besitzen einen hierarchischen Aufbau mit mehreren Dimensionen, der für Fehlerannotationssysteme typisch ist (Díaz-Negrillo/Domínguez-Fernández 2006, Dagneaux/Deness/Granger 1998, Granger 2008: 266). Ein Beispiel für den Aufbau eines Fehlertags in MERLIN im Bereich Präpositionsgebrauch ist in Abbildung 27 dargestellt.

Beispiel: Fehlertag G_Prep_O			
Dimension		Beispiel	Struktur Fehlertag
1	Linguistische Kategorie	Grammatik	G
2	Fehlerspezifikation	Präposition	Prep
3	Zielsprachliche Modifikation (Fehlertyp)	Auslassung	O

Abbildung 27: Aufbau von Fehlertags in MERLIN: Beispiel.

Gebrauch von Partikeln und Idiomen in den Lernertexten erfassen. Auch hierfür liegen entsprechende Annotationstags im Korpus vor (vgl. Wisniewski et al. 2014).

¹⁸³ https://merlin-platform.eu/C_mcorpus.php#anchor6 (Mai 2019).

¹⁸⁴ Das Vorgehen der manuellen, nachträglichen Annotationen im Bereich Präpositionsgebrauch ist im Annotationshandbuch beschrieben (s. Anhang).

Wie in Abbildung 27 deutlich wird, kodieren die Fehlertags in MERLIN 1) die linguistische Kategorie des Fehlers (z. B. *G* für Grammatik), 2) mindestens eine weitere Spezifikation der vorliegenden Abweichung (z. B. *Prep* für Präposition) sowie 3) i.d.R. eine zielsprachliche Modifikation, die den vorliegenden Fehlertyp näher spezifiziert (z. B. Auslassung einer Präposition, hier Kürzel *O*) (Abel et al. 2014: 17). Der Fehlerannotationstag *G_Prep_O* denotiert im Lernerkorpus folglich Abweichungen, die die Auslassung einer Präposition kodieren¹⁸⁵.

Die hierarchische Struktur der Fehlertags erlaubt einen gewissen Grad an Flexibilität bei der Abfrage in MERLIN, da die Struktur der Fehlerannotationstags es ermöglicht, einzelne Ebenen bzw. Dimensionen für die Suche im Lernerkorpus ein- oder auszublenden. So ist es möglich, allgemein nach Fehlern im Bereich der Grammatik zu suchen (Dimension 1), gezielt Präpositionsfehler unterschiedlichen Typs zu extrahieren (Dimension 2) oder spezifische Fehlertypen im Präpositionsgebrauch abzufragen (Dimension 3). Die hierarchische Struktur erlaubt es somit, bestimmte Spezifikationsebenen zu ignorieren bzw. explizit zu berücksichtigen.

6.3 Dateiformat und Korpus-Zugang

Die Korpusdaten liegen im Austauschformat PAULA vor (Christ 1994). Dabei handelt es sich um ein xml-standoff-Dateiformat, d. h. die Annotationen werden, im Gegensatz zu inline-Annotationen, getrennt vom Lernertext gespeichert (Lüdeling et al. 2005)¹⁸⁶. Der Zugriff auf das Lernerkorpus ist derzeit durch das Korpusrecherchesystem ANNIS (Krause/Zeldes 2016, Zeldes et al. 2009) über die Projektplattform (www.merlin-platform.eu) möglich.¹⁸⁷

6.4 Zusammenfassung und Motivation der Korpus-Wahl für die Dissertation

Das Lernerkorpus MERLIN bildet die Datengrundlage für die vorliegende Dissertation. MERLIN ist ein quasi-longitudinales, trilinguales Lernerkorpus, das schriftliche Produktionen von Lernenden der Zielsprachen Deutsch, Italienisch und Tschechisch umfasst. Für die Dissertation ist lediglich das Korpus der Zielsprache Deutsch von Interesse, auf das ich in meiner Dissertation mit der Bezeichnung MERLIN referiere. MERLIN enthält 1033 schriftliche Produktionen (154335 Token) von DaF-Lernenden, die im Rahmen standardisierter Sprachtests zu unterschiedlichen Sprachkompetenzniveaus des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GeR) erhoben und nachträglich bewertet wurden. Das Korpus verfügt über eine Mehr-Ebenen-Architektur mit verschiedenen automatischen und manuellen Annotationen. Zu den zentralen manuellen Annotationen zählen Zielhypothesen (ZH1) sowie

¹⁸⁵ Das Fehlerannotationssystem wurde einheitlich für das gesamte trilinguale Korpus angefertigt (mit einigen sprachspezifischen Fehlertags, vgl. Wisniewski et al. 2014) und liegt daher in englischer Sprache vor.

¹⁸⁶ Um die Konvertierung des ursprünglichen Dateiformats für die Anwendung der einzelnen Annotationswerkzeuge zu gewährleisten, wurde die *SaltNPepper Conversion Suite* (Zipser/Romary 2010) eingesetzt (s. auch Abel et al. 2014: 121).

¹⁸⁷ Das Korpus liegt auf Servern des Instituts für Angewandte Sprachforschung von EURAC in Bozen (Italien). Für eine direkte Korpusrecherche in ANNIS mit erweiterten Exportmöglichkeiten wurde mir für Forschungszwecke ein direkter Zugang zum Korpusrecherchesystem gewährt. Der Datenexport für die empirischen Untersuchungen dieser Dissertation erfolgte somit nicht über die Korpusplattform.

darauf aufbauende Fehlerannotationen. MERLIN ist über das Korpusrecherchesystem ANNIS online frei zugänglich; eine Dokumentation zur Datenerhebung, zum Korpusaufbau und zur Korpusannotation ist ebenfalls online verfügbar¹⁸⁸.

Das Forschungsinteresse der vorliegenden Dissertation liegt in der korpuslinguistischen Ermittlung und Analyse ausgewählter Variablen, die den Gebrauch von Präpositionen und PP durch DaF-Lernende beeinflussen. Insbesondere liegt ein Fokus auf möglichen Schwierigkeiten im Präpositionsgebrauch, die durch den Einfluss einzelner Variablen bedingt werden. Die Wahl des Lernerkorpus MERLIN für das Forschungsvorhaben ist hierbei vor allem durch zwei zentrale Merkmale motiviert:

1) Das Lernerkorpus MERLIN ist nach meinem aktuellem Kenntnisstand das einzige Lernerkorpus der Zielsprache Deutsch, das differenzierte und systematische Fehlerannotationen im Bereich Präpositionsgebrauch enthält und hierfür eine differenzierte Fehlertypologie bereitstellt. Für meine empirischen Untersuchungen kann ich somit auf diese Fehlerannotationen zurückgreifen bzw. im Falle nicht vorhandener Fehlerannotationen entsprechend der Fehlertypologie in MERLIN eine systematische Fehlerkategorisierung gewährleisten.

2) Das Lernerkorpus MERLIN ist nach meinem aktuellem Kenntnisstand das einzige Lernerkorpus der Zielsprache Deutsch, das schriftliche Produktionen von Sprachlernenden unterschiedlicher Kompetenzniveaus enthält, bei denen die Kompetenzniveaus gemäß dem GeR auf Basis textzentrierter Kriterien zuverlässig ermittelt wurden. Die in MERLIN repräsentierten Kompetenzniveaus ermöglichen somit einen zuverlässigen Vergleich des Präpositionsgebrauchs in Abhängigkeit vom Kompetenzniveau sowie eine Skizzierung der kompetenzbedingten Entwicklungen entsprechender Gebrauchsschwierigkeiten.

Weiterhin ist die Korpus-Wahl in der freien Verfügbarkeit von MERLIN sowie der ebenfalls frei verfügbaren und ausführliche Dokumentation zum Korpusaufbau und zur Korpusannotation begründet¹⁸⁹.

7. Fallstudien im Lernerkorpus MERLIN: Überblick

Die empirischen Untersuchungen im Lernerkorpus MERLIN verfolgen das Ziel, die übergeordnete Fragestellung zu Einflussvariablen im Gebrauch von Präpositionen und PP unter besonderer Berücksichtigung von Gebrauchsschwierigkeiten zu beantworten (s. Zusammenfassung Teil I). Die folgende Abbildung (Abb. 28) gibt einen Überblick über die in dieser Dissertation untersuchten Einflussvariablen im Präpositionsgebrauch von DaF-Lernenden. Je untersuchter Einflussvariable sind die jeweiligen Vergleichsgrößen aufgeführt. Zu den untersuchten Einflussvariablen zählen die Erstsprache (L1) der Lernenden, die syntaktische Funktion der PP sowie der semantische Gehalt der Präposition. Der Einfluss der Variable Sprachkompetenz wird in jeder Fallstudie zusätzlich berücksichtigt, da in MERLIN DaF-Lernende mit unterschiedlichen Kompetenzniveaus repräsentiert sind. Nach Granger

¹⁸⁸ Das Korpus sowie die Dokumentation sind verfügbar unter: <http://www.merlin-platform.eu>.

¹⁸⁹ Hierbei handelt es sich jedoch nicht um ein Alleinstellungsmerkmal von MERLIN, da andere Lernerkorpora (z. B. Falko) ebenfalls frei verfügbar sind.

(2015: 12) widmen sich innerhalb der Lernerkorpusforschung vergleichsweise nur wenige Untersuchungen dem Vergleich von Lerneräußerungen auf unterschiedlichen Kompetenzniveaus. Durch die Wahl des Lernerkorpus MERLIN als Datengrundlage in der vorliegenden Dissertation leiste ich folglich einen Beitrag zur Schließung dieser Forschungslücke.

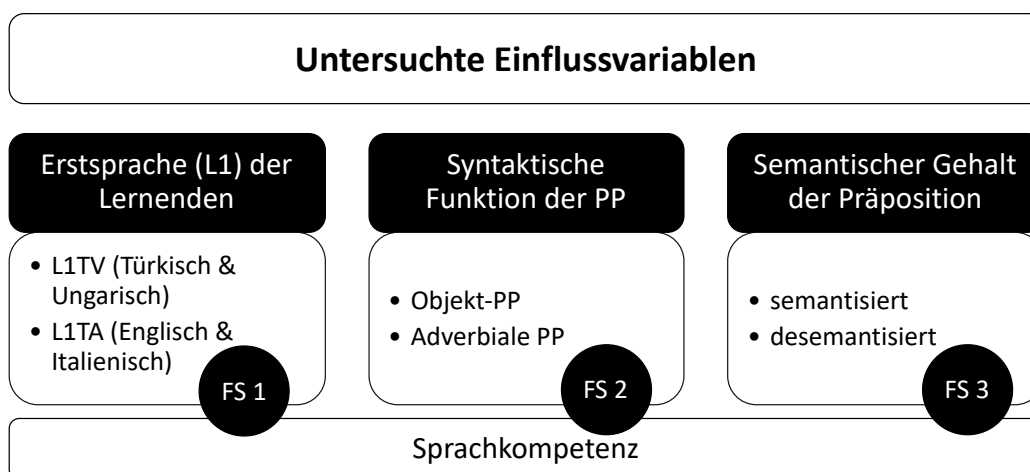


Abbildung 28: Überblick über untersuchte Einflussvariablen im Gebrauch von Präpositionen und PP durch Deutschlernende¹⁹⁰.

Jede Einflussvariable wird in je einer empirischen Fallstudie (FS) im Lernerkorpus MERLIN korpusgestützt untersucht. In jeder Fallstudie werden die entsprechenden Daten mit Rückgriff auf zentrale Analysemethoden der Lernerkorpusforschung (s. Kap. 3.4) analysiert und im Anschluss statistisch ausgewertet. Mittels logistischer Regressionsanalysen in der Statistikumgebung *R* wird der Einfluss der jeweiligen Variable auf die Korrektheit im Präpositionsgebrauch und auf die einzelnen Fehlertypen im Präpositionsgebrauch systematisch untersucht.

Die einzelnen Fallstudien (FS 1, FS 2 und FS 3) sind folgendermaßen aufgebaut: Nach einer Einleitung folgt die Darstellung des Untersuchungsgegenstands (=der jeweiligen Einflussvariable) sowie die Präsentation der zentralen Leitfragen und Hypothesen, die die empirische Untersuchung im Lernerkorpus steuern und sich in die übergeordnete(n) Fragestellung(en) einbetten. Daran anschließend wird die Datengrundlage der jeweiligen Fallstudie bzw. die Datenauswahl und Datenerhebung im Lernerkorpus MERLIN beschrieben bevor auf die Datenaufbereitung eingegangen wird. Die Datenauswertung greift im Anschluss die eingangs formulierten Leitfragen und Hypothesen auf und präsentiert die Ergebnisse der Analysen und statistischen Auswertungen. Zum Schluss erfolgt eine Zusammenfassung und Diskussion der zentralen Ergebnisse. In den folgenden Kapiteln gehe ich näher auf die Analysemethoden (7.1) und statistischen Auswertungen (7.2) der einzelnen Fallstudien ein.

7.1 Analysemethoden

Die ermittelten PP-Kontexte werden in jeder Fallstudie mittels der computerunterstützten Fehleranalyse (computer-aided error analysis, CEA) (Dagneaux/Denness/Granger 1998) sowie

¹⁹⁰ Bei der Variable *Semantischer Gehalt der Präposition* stehen ausschließlich Objekt-PP im Vordergrund. Näheres dazu in Kapitel 10.

der kontrastiven Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (Granger 1996, 2015) ausgewertet. Die genannten Methoden repräsentieren zentrale Analysemethoden in der Lernerkorpusforschung und sind in der Methodologie dieser fest verankert (vgl. Callies 2015). Die Methoden wurden in Kapitel 3.4. ausführlich vorgestellt, in diesem Kapitel gehe ich daher lediglich auf die Anwendung der Analysemethoden in den einzelnen Fallstudien dieser Dissertation ein.

7.1.1 Computerunterstützte Fehleranalyse

Im Rahmen der computerunterstützten Fehleranalyse (computer-aided error analysis, CEA) werden (1) Lerneräußerungen gesammelt, (2) sprachliche Abweichungen (=Fehler) ermittelt, (3) kategorisiert und (4) Erklärungsansätze für die Fehler dargeboten. In den vorliegenden Fallstudien bilden die in MERLIN verfügbaren Fehlerannotationen im Bereich Präpositionsgebrauch die Grundlage für die computerunterstützte Fehleranalyse, d. h. die Auswertung der vorliegenden Abweichungen im Präpositionsgebrauch. Schritt (1) der computerunterstützten Fehleranalyse erfolgt mit der Datenerhebung im Lernerkorpus, die im Einzelnen in jeder Fallstudie (FS) ausführlich beschrieben wird. Schritte (2) und (3) sind in den im Korpus enthaltenen Fehlerannotationen repräsentiert (FS 1) bzw. erfolgen durch sich daran orientierende manuell durchgeführte Fehlerannotationen (in FS 2 und FS 3). Auf Schritt (4) gehe ich jeweils in der Ergebnispräsentation und der anschließenden Ergebnisdiskussion der einzelnen Fallstudien näher ein. An dieser Stelle möchte ich noch erklärende Hinweise zu Schritt (2) und (3) geben.

Die zentrale Voraussetzung für die Fehleridentifikation (Schritt 2) sowie die daran anschließende Fehlerkategorisierung (Schritt 3) ist eine Definition von *Fehler*. Im Rahmen der empirischen Fallstudien definiere ich einen Fehler im Präpositionsgebrauch als eine entsprechende Abweichung zwischen der Original-Lerneräußerung und der Zielhypothese (ZH1). Im Fokus stehen dabei ausschließlich Fehler, die den Gebrauch der Präposition selbst betreffen. Zu diesen Fehlern zählen vorrangig ausgelassene, inkorrekt hinzugefügte und inkorrekt gewählte Präpositionen. Abweichungen, die nicht die Präposition selbst betreffen (sondern z. B. die Kasusreaktion der Präposition), werden für die Datenauswertung nicht berücksichtigt.

Die Kategorisierung vorliegender Fehler im Präpositionsgebrauch (Schritt 3) erfolgt gemäß den Fehlerannotationskategorien im MERLIN-Annotationsschema (Wisniewski et al. 2014: 12). Die folgende Tabelle (Tab. 25) führt die Fehlerannotationskategorien (inkl. der Fehlertags) auf, die in den einzelnen FS berücksichtigt werden¹⁹¹.

¹⁹¹ Der Fehlertyp *pos* (Position einer Präposition) spielt in den einzelnen Fallstudien keine bzw. keine nennenswerte Rolle und wird aus diesem Grund an dieser Stelle nicht aufgeführt. Für ein Beispiel zum Fehlertyp *pos* s. Wisniewski et al. (2014: 12).

Fehlerannotationskategorie (Fehlertag)	Beispiel aus MERLIN
Auslassung einer Präposition (G_Prep_O)	<i>weil ich 9. März eine Prüfung habe.</i>
Hinzufügung einer Präposition (G_Prep_Ad)	<i>Aber man muß auch an der Kultur des Gastlandes kennenzulernen</i>
Wahl einer Präposition (G_Prep_Ch)	<i>Ich freue mich um Ihrer Angebot.</i>

Tabelle 25: Fehlerannotationskategorien im Präpositionsgebrauch in MERLIN: Beispiele.

Für die Ermittlung inkorrektter PP-Kontexte werden in Anlehnung an die *Target-like Use Analysis* (vgl. Pica 1984, Ellis/Barkhuizen 2005: 74) sowohl PP-Kontexte berücksichtigt, in denen zielsprachlich (d. h. auf Ebene der ZH1) eine PP gefordert ist, als auch solche Kontexte, in denen der/die Lernende eine Präposition hinzugefügt hat, die zielsprachlich nicht gefordert ist (d. h. auf Ebene der ZH1 getilgt wurde). Diese PP-Kontexte werden jeweils bei der Abfrage im Lernerkorpus berücksichtigt. Die Berücksichtigung beider Kontexte (zielsprachlich gefordert/zielsprachlich nicht gefordert) liefert nicht nur einen Hinweis darauf, wie erfolgreich Lernende eine Präposition gelernt haben und diese in zielsprachlich geforderten syntaktischen Kontexten anwenden können. Sie liefert auch wertvolle Hinweise darauf, ob die Lernenden die Produktion dieser Präposition auch kontrollieren können, folglich ebenfalls die syntaktische Distribution der jeweiligen Präposition erlernt haben (vgl. Pica 1984: 74)¹⁹².

Im Rahmen der computerunterstützten Fehleranalyse wird nicht zwischen Performanz- und Kompetenzfehlern (vgl. Ellis/Barkhuizen 2005: 62) unterschieden, da dies auf Basis der Datengrundlage nicht möglich ist. Ebenfalls erfolgt keine Trennung zwischen ‚versteckten‘ Fehlern (covert errors) und an der sprachlichen Oberfläche evidenten Fehlern (overt errors, vgl. Callies 2015: 41) (s. hierzu Kap. 3.4.1). Die Ergebnisse der computerunterstützten Fehleranalyse fließen in die kontrastive Interlanguage-Analyse ein, auf die ich im Folgenden eingehe.

7.2.2 Kontrastive Interlanguage-Analyse

Bei der kontrastiven Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (Granger 1996, 2015) werden Frequenzen ausgewählter Kategorien in verschiedenen Korpora verglichen (Lüdeling/Walter 2010: 319). In den Fallstudien dieser Dissertation erfolgt ein Vergleich qualitativer Unterschiede (Korrektheit und Fehlertypen) im Präpositionsgebrauch unter Berücksichtigung der einzelnen Einflussvariablen (Erstsprache der Lernenden, Sprachkompetenz, Syntaktische Funktion der PP und Semantischer Gehalt der Präposition).

In allen Fallstudien werden a) unterschiedliche Ausprägungen der Einflussvariablen und b) Lernende mit unterschiedlichen Kompetenzniveaus (GeR-Gesamtniveaus) kontrastiert. Die Vergleiche tragen dazu bei, Einflussvariablen im Sprachgebrauch zu ermitteln. Mit der kontrastiven Interlanguage-Analyse untersuche ich in den einzelnen Fallstudien jeweils, welche Ergebnisse je Einflussvariable und GeR-Gesamtniveau der Lernenden für die Korrektheit und für die Fehlertypen im Präpositionsgebrauch ermittelt werden können und

¹⁹² Pica (1984) diskutiert die *Target-like Use Analysis* im Kontext von Morphemstudien zur L2 Englisch, jedoch lässt sich die Analyse auch auf sprachliche Einheiten übertragen, die keinen Morphemstatus besitzen.

inwieweit diese wiederum durch das GeR-Gesamtniveau und/oder die ausgewählte Einflussvariable beeinflusst werden. Die folgende Tabelle (Tab. 26) gibt eine Zusammenfassung der Analysemethoden, die in allen drei Fallstudien angewandt werden sowie die damit verbundenen Ziele.

Analysemethode	Ziel
Computerunterstützte Fehleranalyse	Ermittlung der Abweichungen (bzw. Korrektheit) im Präpositionsgebrauch (in zielsprachlich geforderten und zielsprachlich nicht geforderten PP-Kontexten). Die Fehlerkategorisierung erfolgt auf Basis der Fehlerkategorien im MERLIN-Annotationschema (Wisniewski et al. 2014: 12).
Kontrastive Interlanguage-Analyse	Kontrastiver Vergleich der Ergebnisse der computerunterstützten Fehleranalyse je GeR-Gesamtniveau der Lernenden und ausgewählter Einflussvariable (Erstsprache der Lernenden, Sprachkompetenz, Syntaktische Funktion der PP und Semantischer Gehalt der Präposition).

Tabelle 26: Analysemethoden der Fallstudien: Überblick.

7.2 Hinweise zur statistischen Auswertung

In allen Fallstudien der Dissertation erfolgt die statistische Auswertung mittels einer (logistischen) Regressionsanalyse. In diesem Kapitel möchte ich daher näher auf die Regressionsanalyse eingehen, ihren Einsatz in der vorliegenden Dissertation motivieren und die zentralen Parameter der Regressionsanalyse in den einzelnen empirischen Fallstudien explizieren.

7.2.1 Regressionsanalyse

Eine Regressionsanalyse „dient der Analyse von Beziehungen zwischen einer abhängigen Variablen und einer oder mehreren unabhängigen Variablen“ (Backhaus et al. 2016: 63). Eine Regressionsanalyse wird vor allem eingesetzt für die quantitative Beschreibung und Erklärung von Zusammenhängen sowie der Schätzung bzw. Prognose von Werten der abhängigen Variable (Backhaus et al. 2016: 63) und ist daher angemessen für das Forschungsvorhaben der vorliegenden Dissertation, in der der Einfluss ausgewählter (unabhängiger) Variablen auf den Gebrauch von Präpositionen und PP (=abhängige Variable bzw. Zielvariable) untersucht wird. Eine logistische Regression ist eine spezielle Variante einer Regressionsanalyse, bei der die abhängige Variable eine kategoriale Variable darstellt. Die Ausprägungen der kategorialen Variable repräsentieren die Alternativen. Bei einer logistischen Regression gilt die abhängige Variable als eine Zufallsvariable und es werden Wahrscheinlichkeiten für die einzelnen Alternativen der abhängigen Variable prognostiziert (vgl. Backhaus et al. 2016: 284). Liegen zwei Alternativen vor, spricht man auch von einer binären logistischen Regression (vgl. Backhaus et al. 2016: 284).

In der Lernerkorpusforschung wird die Regressionsanalyse typischerweise im Kontext hypothesenüberprüfender Untersuchungen angewandt (Gries 2015: 16), binäre logistische

Regressionsanalysen sind nach Gries (2013: 300) „now common“ in korpuslinguistischen Studien. Die Regressionsanalyse ist ein Ansatz der statistischen Auswertung, der Untersuchungen zum Effekt verschiedener Variablen auf den Gebrauch linguistischer Strukturen ermöglicht (Gries 2015: 161)¹⁹³. Somit stellt die Regressionsanalyse in der Lernerkorpusforschung einen mächtigen Ansatz der Datenauswertung dar. Nach Gries (2015: 166) bietet eine Regressionsanalyse im Kontext der Lernerkorpusforschung u.a. folgende Vorteile gegenüber alternativen Ansätzen der statistischen Auswertung¹⁹⁴:

- Parallele Berücksichtigung mehrerer (unabhängiger) Variablen
- Untersuchung von Interaktionen mehrerer (unabhängiger) Variablen

Mit einer Regressionsanalyse ist es möglich, den Einfluss bzw. Effekt unterschiedlicher Variablen parallel zu berechnen. Die parallele Berücksichtigung mehrerer Variablen spiegelt auch gut die multifaktorielle Natur linguistischer Phänomene wider, denn „there is always more than one cause for any given effect“ (Gries 2013: 300). Eine Interaktion zwischen mehreren (unabhängigen) Variablen wird definiert als die nicht-additive Verknüpfung der Effekte, die die (unabhängigen) Variablen einzeln aufweisen. Somit ist es mithilfe einer Regressionsanalyse möglich zu bestimmen, ob eine (unabhängige) Variable einen Effekt hat auf den Zusammenhang zwischen einer anderen (unabhängigen) Variable und der (abhängigen) Zielvariable (Gries 2013: 300, Backhaus et al. 2016: 101). Interaktionen mehrerer (unabhängiger) Variablen liefern somit wertvolle Hinweise darauf, inwieweit die Effekte einzelner Variablen sich gegenseitig beeinflussen.

Eine Regressionsgleichung (auch: Regressionsmodell) modelliert eine (abhängige) Zielvariable als eine Funktion (\sim) von einer (o. mehrerer) (unabhängiger/n) Variablen und ggf. deren Interaktionen (Gries 2015: 166). Der Output des Regressionsmodells liefert je unabhängiger Variable (bzw. Interaktion) unterschiedliche Werte, die in Bezug auf das Regressionsmodell interpretiert werden müssen (Gries 2013: 301). Als Beispiel sei im Folgenden ein Regressionsmodell und sein Output aus Gries (2013: 301) dargestellt:

GENITIVE \sim ANIM_POSSOR + SPEAKER + ANIM_POSSOR:SPEAKER

Im Beispiel geht es um Einflussvariablen auf den Gebrauch des Genitiv-s (GENITIVE) bei Englischlernenden (Näheres s. Gries 2013: 290ff.), zu dem aufbereitete Daten aus einem Lernerkorpus und einem muttersprachlichen Vergleichskorpus herangezogen werden. Der Gebrauch des Genitiv-s stellt die Zielvariable dar. Als unabhängige Variablen werden hier die Belebtheit des Besitzenden (ANIM_POSSOR) sowie der/die Sprecher/-in (SPEAKER), d. h. Produzent/-in des Genitiv-s angesetzt. Zusätzlich werden Interaktionen zwischen diesen Variablen berücksichtigt (durch Doppelpunkt markiert)¹⁹⁵. Die Ausprägungen der kategorialen

¹⁹³ Siehe z. B. Gries/Wulff (2013) für die Anwendung einer Regressionsanalyse in der Lernerkorpusforschung.

¹⁹⁴ Gries (2015: 160f.) nennt hier z. B. die Berechnung von Assoziationsmaßen oder verschiedene Arten von Signifikanztests.

¹⁹⁵ Das Plus (+) markiert eine Aufzählung einzelner Variablen (und deren Einzeleffekte), der Doppelpunkt (:) markiert eine Interaktion zwischen den entsprechenden Variablen. Mit einem Sternchen (*) können beide Operationen parallel ausgedrückt werden: GENITIVE \sim ANIM_POSSOR*SPEAKER würde dann die Einzeleffekte entsprechender Variablen sowie deren

Variablen sind jeweils belebt vs. unbelebt (engl. animate vs. inanimate) (bei ANIM_POSSOR) sowie MuttersprachlerInnen/Lernende (engl. native vs. learner) (bei SPEAKER). Der Output zum obigen Regressionsmodell liefert die in Abbildung 29 angegebenen Werte, die Nummerierungen habe ich zur besseren Orientierung nachträglich eingefügt.

		Coef ¹⁹⁶	S.E. ¹⁹⁷	Wald z ¹⁹⁸	Pr (> z)
0	Intercept	0,3697	0,2109	1,75	0,0796
1	ANIM_POSSOR=inanimate	- 2,7836	0,3295	- 8,45	<0,0001
2	SPEAKER=native	1,4826	0,3418	4,34	<0,0001
3	ANIM_POSSOR=inanimate *SPEAKER=native	- 0,0955	0,4574	0,21	0,8347

Abbildung 29: Output eines Regressionsmodells mit zwei unabhängigen Variablen und einer Interaktion zwischen den Variablen: Beispiel aus Gries (2013: 301).

Ich möchte an dieser Stelle auf die zwei wichtigsten Werte des Modell-Outputs eingehen: Den Koeffizienten (Coef) und den p-Wert (Pr (>|z|)¹⁹⁹ (Abb. 29). Der Koeffizient wird gemeinsam mit seinem Standardfehler (S.E.) angegeben und zeigt den Effekt der entsprechenden Variable an. Negative bzw. positive Koeffizienten spiegeln dabei die Richtung des Effekts der entsprechenden Variable wider, d. h. Abnahme bzw. Zunahme der prognostizierten Wahrscheinlichkeit der abhängigen Zielvariable. Der p-Wert gibt jeweils an, ob ein signifikanter Effekt der entsprechenden unabhängigen Variable auf die abhängige Zielvariable vorliegt (Gries 2013: 301). Wie man den Werten im Modell-Output (Abb. 29) entnehmen kann, können für das Beispiel-Regressionsmodell signifikante Einzeleffekte der Variablen ANIM_POSSOR und SPEAKER ermittelt werden (s. p-Werte bei 1 und 2 in Abb. 29)²⁰⁰. Der negative Koeffizient bei ANIM_POSSOR=inanimate (1) zeigt dabei an, dass die Ausprägung *inanimate* die Wahrscheinlichkeit für das Genitiv-s (die Zielvariable) verringert, während der positive Koeffizient bei SPEAKER=native (2) anzeigt, dass die Ausprägung *native* die Wahrscheinlichkeit für die Zielvariable erhöht. Weiterhin kann man feststellen, dass keine Interaktion zwischen den beiden unabhängigen Variablen im Modell vorliegt (s. p-Wert bei 3)²⁰¹. Wenn keine Interaktion zwischen den im Regressionsmodell enthaltenen Variablen festgestellt wird, gilt es, das Regressionsmodell entsprechend anzupassen: Das bedeutet, dass man die Interaktion zwischen den unabhängigen Variablen aus dem Regressionsmodell entfernt und ausschließlich die Einzeleffekte der unabhängigen Variablen berücksichtigt. Das obige Regressionsmodell wäre dann folgendermaßen zu modifizieren (vgl. Gries 2013: 302):

GENITIVE ~ ANIM_POSSOR + SPEAKER

Interaktion kodieren und wäre damit inhaltlich identisch zum im obigen Beispiel dargestellten Regressionsmodell.

¹⁹⁶ In R als *Estimate* angegeben.

¹⁹⁷ In R als *Std. Error* angegeben.

¹⁹⁸ In R als *z value* angegeben.

¹⁹⁹ Zur Erklärung aller Werte s. z. B. Gries (2013: 302).

²⁰⁰ Als Signifikanzniveau wird hier 5% angenommen.

²⁰¹ Die Werte in der Zeile *Intercept* (0) sind für die Interpretation des Modell-Outputs nicht relevant (vgl. auch Gries 2013: 301). Aus diesem Grund gehe ich nicht weiter auf diese Zeile ein.

Die Ergebnisse dieses Modells wären wieder entsprechend den obigen Ausführungen zu interpretieren und das Modell ggf. weiter anzupassen.

In einem Regressionsmodell können Effekte einzelner unabhängiger Variablen auf eine Zielvariable berechnet bzw. prognostiziert werden. Auf diese Weise können z. B. gezielt Erkenntnisse zu Einflussfaktoren im Sprachgebrauch ermittelt werden. Feste Faktoren, deren Effekt ermittelt werden soll, werden auch *fixed effects* genannt. Weiterhin können Zufallsfaktoren (*random effects*) in ein Regressionsmodell integriert werden. Zufallsfaktoren sind Variablen, deren Einfluss im Modell zwar berücksichtigt, jedoch ‚herausgerechnet‘ wird, um die Ergebnisse des Regressionsmodells über diese Variablen hinweg generalisieren zu können (Larson-Hall 2010: 40). Im Kontext von Untersuchungen zum Sprachgebrauch durch Lernende ist es auf diese Weise möglich, z. B. den Einfluss einzelner Lernender herauszurechnen, um die Ergebnisse des Regressionsmodells über die im Datensatz enthaltenen Lernenden hinaus interpretieren zu können. Nach Wolfer (2017: 145) kann die Berücksichtigung von Zufallsfaktoren dazu führen, dass „Effekte der festen Faktoren sauberer hervortreten“. Regressionsmodelle, die feste Faktoren und Zufallsfaktoren kombinieren, nennt man auch *gemischte Modelle* (Wolfer 2017: 145) (engl. *mixed-effects models*, s. Larson-Hall 2010: 40). In der vorliegenden Dissertation werden in jeder Fallstudie gemischte Modelle berechnet und einzelne Lernende als Zufallsfaktoren behandelt. Dies ist auch insofern notwendig, als in meinen Datensätzen jeweils i.d.R. mehrere Datenpunkte (d. h. Treffer im Lernerkorpus) einer/m einzelnen Lernenden zugewiesen werden können. Der/die einzelne Lernende hat dann einen Einfluss auf die Varianz der sprachlichen Daten im Datensatz und dieser Einfluss muss in der statistischen Auswertung berücksichtigt werden (vgl. Tagliamonte/Baayen 2012: 143)²⁰². Dies ist über ein gemischtes Modell möglich.

Das Regressionsmodell kann mit der Berechnung des R^2 -Werts evaluiert werden (Gries 2013: 302). Der R^2 -Wert gibt an, wie gut die vorliegenden Daten auf Basis des Regressionsmodells erklärt werden können. Der R^2 -Wert kann zwischen 0 und 1 liegen – dabei gilt nach Gries (2013: 302): Je höher der Wert, desto besser. Jedoch möchte ich an dieser Stelle zusätzlich darauf hinweisen, dass der R^2 -Wert immer im Kontext des jeweiligen Regressionsmodells bzw. der jeweiligen zugrundeliegenden Untersuchung interpretiert werden muss. Ein niedriger R^2 -Wert ist dabei nicht grundsätzlich als schlecht einzustufen (Minitab Blog 2013).

Um die Ergebnisse des Modells und die Effekte einzelner Variablen adäquat zu interpretieren, sind Visualisierungen sehr hilfreich. Bei komplexen Regressionsmodellen sind diese sogar essenziell für das Verständnis des entsprechenden Modell-Outputs (Gries 2013: 302ff., Wolfer/Hansen-Morath 2018: 242). Dabei gibt es für Visualisierungen von Regressionsmodellen unterschiedliche Möglichkeiten (s. Gries 2013: 302ff.). Auf die Visualisierungen der Modell-Outputs in den Fallstudien dieser Dissertation wird unten näher eingegangen.

²⁰² Aus diesem Grund ist z. B. die Anwendung eines Chi²-Tests (Meindl 2011: 162) bei derartigen Datensätzen nicht möglich, da der Chi²-Test unter der Grundannahme operiert, dass die einzelnen Datenpunkte unabhängig voneinander sind, d.h. jeweils unterschiedlichen Lernenden zugewiesen werden können (s. auch Larson-Hall 2010: 206).

7.2.2 Spezifische Anmerkungen zur Regressionsanalyse in der vorliegenden Dissertation

Der empirische Teil der vorliegenden Dissertation gliedert sich in drei Fallstudien. In jeder Fallstudie werden die entsprechenden Daten aus dem Lernerkorpus MERLIN im Hinblick auf die jeweiligen Fragestellungen tabellarisch aufbereitet²⁰³ und mittels einer Regressionsanalyse statistisch ausgewertet²⁰⁴. Hierfür wählte ich die Statistik-Umgebung *R* (R Core Team 2018) bzw. die Benutzeroberfläche *R-Studio* (RStudio, Inc. 2018) (Version 1.1.456). Die Statistik-Umgebung *R* ist unter freier Lizenz verfügbar und wird mittlerweile vermehrt für statistische Berechnungen im Kontext wissenschaftlicher Arbeiten genutzt (Tippmann 2014).

In jeder Fallstudie der vorliegenden Dissertation werden mit der *glmer* Funktion aus dem Packet *lme4* in *R* (Bates et al. 2015) gemischte Modelle (*generalized linear mixed models*) berechnet. Innerhalb der Modelle wird der *family*-Parameter jeweils auf *binomial* („family=binomial“) gesetzt, da jeweils eine logistische Regression mit kategorialen Variablen vorliegt (s. Wolfer 2017: 160, Fußnote). Der/die einzelne Lernende wird jeweils als Zufallsfaktor (*random factor*) in das jeweilige Regressionsmodell integriert²⁰⁵. Weiterhin wird für die Regressionsmodelle jeweils der *bobyqa-optimizer* festgelegt.

In jeder Fallstudie werden Effekte einzelner (unabhängige) Einflussvariablen auf den Präpositionsgebrauch von DaF-Lernenden untersucht, wobei die Korrektheit und die Fehlertypen im Präpositionsgebrauch im Vordergrund stehen. Dabei gibt es in einzelnen Regressionsmodellen (unabhängige) kategoriale Variablen mit zwei Ausprägungen sowie eine kategoriale Variable, das GeR-Gesamtniveau, welches in einigen Regressionsmodellen drei (bzw. vier) Ausprägungen aufweist (A2, B1, B2 und ggf. C1). Wie weiter oben in Abbildung 29 dargestellt, liefert der Output eines Regressionsmodells je (unabhängiger) Variable eine Zeile mit u.a. dem Koeffizient- bzw. Estimate-Wert und p-Wert. Dabei ist jeweils eine Ausprägung der entsprechenden Variable angegeben, die andere(n) Ausprägung(en) dient/(dienen) als Referenzwert. Das bedeutet, dass die im Modell-Output angegebenen Werte in Bezug auf diesen Referenzwert zu interpretieren sind. Bei kategorialen Variablen mit zwei Ausprägungen ist der Referenzwert immer die jeweils andere Ausprägung, im Falle einer kategorialen Variable mit mehr als zwei Ausprägungen gibt es jedoch mindestens zwei (bzw. drei) Referenzwerte und entsprechend gibt es mindestens zwei (bzw. drei) Zeilen im entsprechendem Modell-Output. Ähnliches gilt auch für die Berechnung der Interaktion zwischen unterschiedlichen Variablen: Auch hier werden im Modell-Output für kategoriale Variablen mit mehr als zwei Ausprägungen mehrere Zeilen ausgegeben.

²⁰³ Die entsprechenden Tabellen bilden den jeweiligen Datensatz ab und enthalten u.a. einzelne Spalten zu Korrektheit, Fehlertyp, Identifikationsnummer, Erstsprache etc. Die entsprechenden Werte (=Datenpunkte) sind jeweils in einzelnen Zeilen der Tabelle enthalten. Eine solche Tabelle bildet den Input für die Berechnung von Regressionsmodellen in *R*.

²⁰⁴ An dieser Stelle möchte ich mich sehr herzlich bei Sascha Wolfer bedanken, der mich in Bezug auf die statistischen Auswertungen beraten hat.

²⁰⁵ Das entsprechende Kürzel im Regressionsmodell in *R* lautet (1|Lerner). Die Wahl der maskulinen Form *Lerner* meint an dieser Stelle nicht, dass andersgeschlechtige Lernende nicht eingeschlossen sind. Die Form *Lerner* im Regressionsmodell referiert lediglich auf die entsprechende Spalte der Tabelle, die als Input für das jeweilige Regressionsmodell dient. Die Spalte *Lerner* enthält die unterschiedlichen Identifikationsnummern der Lernertexte im jeweils vorliegenden Datensatz.

Zur Illustration möchte ich an dieser Stelle gerne das oben erwähnte Beispiel zum Gebrauch des Genitiv-s aufgreifen. Der Gebrauch des Genitiv-s repräsentiert die (abhängige) Zielvariable. Angenommen, es soll der Einfluss des/der Sprechers/-in sowie der Sprachkompetenz der Lernenden ermittelt werden. Die (unabhängige) Variable *Sprecher/-in* (SPEAKER) hat zwei Ausprägungen (native bzw. non-native), die (unabhängige) Variable *Sprachkompetenz* (GeR) hat drei Ausprägungen: A2, B1 und B2. Im Output des entsprechenden Regressionsmodells gilt nun eine Ausprägung der Sprachkompetenz (A2) als Referenzwert, die Effekte der anderen beiden Ausprägungen (B1 und B2) werden nur in Bezug auf diesen Referenzwert angegeben, vgl. die folgende Abbildung (Abb. 30) zur Illustration eines solchen Modell-Outputs²⁰⁶.

In der Abbildung (Abb. 30) sieht man, dass der *allgemeine* Effekt einer kategorialen Variable mit mehr als zwei Ausprägungen (hier: GeR) auf die Zielvariable sowie *die allgemeine* Interaktion zwischen einer Variable mit zwei Ausprägungen (hier: SPEAKER) und einer Variable mit mehr als zwei Ausprägungen (hier: GeR) nicht im Modell-Output ausgegeben wird. Interaktion und Effekt werden jeweils lediglich in Bezug auf den Referenzwert (hier: A2) ausgegeben.

		Estimate	Std. Error	z value	Pr (> z)
0	(Intercept)				
1	SPEAKER=ative				
2	GeR_B1				
3	GeR_B2				
4	SPEAKER=ative: GeR_B1				
5	SPEAKER=ative: GeR_B2				

Abbildung 30: Output eines Regressionsmodells mit zwei kategorialen Variablen (SPEAKER und GeR) mit unterschiedlicher Anzahl an Ausprägungen: Beispiel.

Zur Ermittlung der *allgemeinen* Effekte von Variablen mit mehr als zwei Ausprägungen (hier: GeR) sowie zur Ermittlung der *allgemeinen* Interaktion zwischen einer Variable mit zwei Ausprägungen (hier: SPEAKER) und der Variable mit mehr als zwei Ausprägungen (hier: GeR) habe ich in meinen empirischen Untersuchungen die *Anova()*-Funktion des Pakets *Car* (Fox/Weisberg 2011) in *R* angewendet. Die Anwendung der *Anova()*-Funktion auf das entsprechende Regressionsmodell liefert u.a. einzelne *p*-Werte²⁰⁷ (Spalte *Pr (>Chisq)*) je *allgemeinem* Effekt einer (unabhängiger) Variable bzw. je *allgemeiner* Interaktion zwischen den Variablen, unabhängig von der Anzahl der Ausprägungen dieser Variablen (vgl. Abb. 31 zur Illustration).

²⁰⁶ Die Benennungen im Modell-Output sind *R* entnommen. Auf die Angabe entsprechender Werte in den einzelnen Zellen der Tabelle wird an dieser Stelle verzichtet, da diese für Illustrationszwecke nicht relevant sind.

²⁰⁷ Der zugrundeliegende Signifikanztest ist der Wald Signifikanztest (Typ II).

	Response: Zielvariable	Chisq	Df ²⁰⁸	Pr (>Chisq)
1	SPEAKER			
2	GeR			
3	SPEAKER:GeR			

Abbildung 31: Output der Anova()-Funktion in R zu einem Regressionsmodell mit zwei unabhängigen kategorialen Variablen: Beispiel.

Durch die Anwendung der Anova()-Funktion erhält man Aussagen darüber, ob die im Regressionsmodell enthaltenen (unabhängigen) kategorialen Variablen mit mehr als zwei Ausprägungen (hier: GeR) einen allgemeinen (signifikanten) Effekt auf die Zielvariable haben und ob im Modell eine allgemeine Interaktion zwischen einer Variable mit zwei Ausprägungen (hier: SPEAKER) und einer Variable mit mehr als zwei Ausprägungen (hier: GeR) vorliegt. Die Richtung des Effekts lässt sich auf Basis des Outputs der Anova()-Funktion jedoch nicht ableiten, hierfür ist eine Visualisierung des entsprechenden Regressionsmodells notwendig.

Zur Evaluation der Regressionsmodelle wurde mit der *rsquared*-Funktion des Pakets *piecewiseSEM* (Lefcheck 2016) der R^2 -Wert ermittelt. Die Visualisierungen der Regressionsmodelle wurden mit dem Paket *effects* (Fox 2003, Fox/Hong 2009) erzeugt, s. auch Wolfer/Hansen-Morath (2018) zu diesem Paket und weiteren Visualisierungsoptionen in R.

In der vorliegenden Dissertation gelten in Bezug auf die statistischen Auswertungen und ihre Darstellungen folgende Konventionen:

- Die Anova()-Funktion wird nur dann angewandt, wenn kategoriale Variablen mit mehr als zwei Ausprägungen vorliegen. Liegt eine Variable mit zwei Ausprägungen vor, so findet die Anova()-Funktion keine Anwendung und es werden lediglich die Werte im Output des entsprechenden Regressionsmodells präsentiert.
- Das Ergebnis der Anova()-Funktion wird jeweils in verkürzter Form angegeben, wobei lediglich der entsprechende p-Wert zitiert wird. Sofern nicht anders angegeben, handelt es sich an Stellen, an denen lediglich ein p-Wert erwähnt wird, um den p-Wert der entsprechenden Anova()-Funktion. Es wird von einem Signifikanzniveau von 5% ausgegangen. Der vollständige Output der jeweiligen Anova()-Funktion findet sich im Anhang²⁰⁹.
- Beim Output einzelner Regressionsmodelle werden in der Arbeit lediglich die zentralen Werte des Modell-Outputs angegeben (vgl. Abb. 29). Es wird von einem Signifikanzniveau von 5% ausgegangen. Die vollständigen Modell-Outputs (so wie in R-Studio ausgegeben) sind auf Anfrage erhältlich.
- Bei einer vorliegenden Interaktion zwischen Variablen wird auf eine Darstellung der zentralen Werte des entsprechenden Modell-Outputs verzichtet, da diese für sich genommen keine gültige Interpretation zulassen (vgl. Gries 2013: 304). Die zentralen Werte des Modell-Outputs solcher Modelle sind im Anhang zu finden.

²⁰⁸ Df gibt die entsprechenden Freiheitsgrade an.

²⁰⁹ Im Anhang sind die einzelnen Outputs nach Fallstudie sortiert.

- Das Ergebnis der rsquared()-Funktion wird in verkürzter Form wiedergegeben, wobei lediglich der R²-Wert zitiert wird. Der vollständige Output der jeweiligen rsquared()-Funktion findet sich im Anhang.
- Allgemeine Effekte bzw. Interaktionen, die mittels der Anova()-Funktion berechnet werden, (s.o.) sowie der R²-Wert werden lediglich für die finalen Regressionsmodelle angegeben.

8. Fallstudie 1: Erstsprache der Lernenden

8.1 Einleitung

In dieser Fallstudie im Lernerkorpus MERLIN wird die Variable der *Erstsprache (L1)* und ihr Einfluss auf den Gebrauch von Präpositionen und PP durch DaF-Lernende untersucht. Schwierigkeiten im Präpositionsgebrauch zeigen sich bei allen Deutschlernenden, werden in der Forschung jedoch teilweise auch spezifischen Einflüssen der L1 zugeschrieben. Innerhalb der Z/FSE-Forschung ist weitgehend anerkannt, dass die L1 einen Einfluss auf den Erwerb und Gebrauch bestimmter grammatischer Strukturen der Zielsprache (L2) hat (Brdar-Szabó 2010a: 524, Uhlisch 1995: 226). Aktuelle gebrauchsbasierte Ansätze zum Zweit- und Fremdspracherwerb (s. Ellis/Wulff 2015, Mitchell/Myles/Marsden 2013, s. Kap. 2) postulieren einen Einfluss bereits erlernter sprachlicher Strukturen (z. B. der L1) und Form-Funktionsverbindungen auf die Verarbeitung und den Erwerb sprachlicher Strukturen einer L2 (vgl. z. B. Ellis/Wulff 2015: 82). Der Einfluss der L1 wurde ebenfalls für den Erwerb und Gebrauch von deutschen Präpositionen und PP beobachtet²¹⁰. Hier liegen vorwiegend Untersuchungen zu Lernenden mit L1 Türkisch vor (vgl. Kap. 4.1.2). Die vorliegende Fallstudie möchte sich neben dem Türkischen als L1 weiteren L1 von Lernenden widmen und somit umfassende Erkenntnisse zum Einfluss der Erstsprache auf den Gebrauch von Präpositionen und PP durch DaF-Lernende liefern. Neben dem Türkischen werden im Lernerkorpus MERLIN Lernende mit den Erstsprachen Ungarisch, Italienisch und Englisch ausgewertet. Vor allem für das Ungarische fehlen nach Brdar-Szabó (2010b: 732) „empirische Untersuchungen im Bereich des Deutschen aus lernersprachlicher Sicht.“

8.2 Untersuchungsgegenstand und Leitfragen

Ziel dieser Fallstudie ist es zu ermitteln, ob bzw. welchen Einfluss die L1 der DaF-Lernenden auf den Gebrauch von deutschen Präpositionen und PP hat, wobei die Korrektheit und Fehlertypen im Präpositionsgebrauch fokussiert werden. Wie oben erwähnt, gehen aktuelle gebrauchsbasierte Spracherwerbtheorien davon aus, dass bereits erlernte Sprachen²¹¹ die Verarbeitung und Produktion sprachlicher Strukturen der L2 beeinflussen. Das Phänomen der

²¹⁰ Einige Untersuchungen zum Erwerb und Gebrauch von Präpositionen konnten keinen spezifischen Einfluss der L1 feststellen (vgl. Steinmüller 1992, Fries 1988, Gibson/Hufeisen/Libben 2001 und Turgay 2010a) (s. Kap. 4.1.2).

²¹¹ Zu den bereits erworbenen Sprachen zählen neben der L1 auch Zweit- oder Fremdsprachen. Ich möchte mich an dieser Stelle jedoch ausschließlich auf die Erstsprache der Lernenden konzentrieren.

„learned attention“ (Ellis/Wulff 2015: 82) beschreibt die (unbewusste) Aufmerksamkeit von Sprachlernenden auf bereits (z. B. in der L1) erworbene und verfestigte Form-Funktionsverbindungen. Die „learned attention“ kann auf diese Weise zur ‚Überschattung‘ anderer, in der zu erlernenden Zielsprache jedoch relevanter sprachlicher Merkmale („cues“), führen, denn „redundant cues are overshadowed because the learners’ L1 experience leads them to look elsewhere for the cues to interpretation“ (Ellis/Wulff 2015: 83). Somit kann die Aufmerksamkeit auf gewisse sprachliche Strukturen der Zielsprache bedingt durch das sprachliche Wissen aus der L1 in gewisser Weise blockiert werden²¹² (s. auch Ausführungen in Kap. 2). Basierend auf diesen Erkenntnissen kann man daher annehmen, dass der Erwerb und Gebrauch sprachlicher Strukturen von den Gemeinsamkeiten und Unterschieden zwischen L1 und L2 beeinflusst wird (Ellis 2006b: 175). Lütke (2011: 111) fasst in einem kurzen Forschungsüberblick zu diesem Thema das Folgende zusammen:

Untersuchungen zum Gebrauch lokaler Präpositionen von kindlichen Lernenden der L2 Deutsch lassen vermuten, dass der Erwerb präpositionaler Wendungen für Lernende, die eine typologisch entfernte L1 wie das Türkische sprechen, deutlich schwieriger ist.

Für die typologische Nähe bzw. Distanz einer L1 zur Zielsprache Deutsch wird daher ein Einfluss auf den L2-Erwerb angenommen. Dies gilt nicht nur für den L2-Erwerb durch Kinder, wie die Ausführungen in Kapitel 4.1.2 verdeutlichen. Blumenthal-Dramé/Kortmann (2013: 293) erwähnen, dass insbesondere solche sprachlichen Merkmale der L2 den Lernenden Schwierigkeiten bereiten, „in denen die Fremdsprache aus typologischer Sicht ungewöhnlicher ist“ als die L1. Daraus ergibt sich die zentrale Motivation für die Auswahl der zu untersuchenden L1 in der vorliegenden Fallstudie: Die Fallstudie widmet sich dem Vergleich des Präpositionsgebrauchs bei DaF-Lernenden mit typologisch unterschiedlichen L1.

Zu den *typologisch verschiedenen* (TV) Sprachen zählen hier exemplarisch das Türkische und das Ungarische. In beiden Sprachen werden Relationen, die im Deutschen durch Präpositionen bzw. PP realisiert werden, vorrangig mithilfe von Kasussuffixen ausgedrückt (vgl. Schroeder/Şişmek 2010: 721, Brdar-Szabó 2010b: 733). Zu den *typologisch ähnlichen* (TA) Sprachen zähle ich hier exemplarisch die Sprachen Englisch und Italienisch. Beide Sprachen verfügen, ähnlich zum Deutschen, über Präpositionen und PP, die in unterschiedlichen syntaktischen Kontexten realisiert werden können (Näheres zu Präpositionen und PP aus kontrastiver Perspektive s. Kap. 1.3). In der vorliegenden Fallstudie werden die L1 Türkisch und Ungarisch den L1 Englisch und Italienisch gegenübergestellt. Die folgende Tabelle (Tab. 27) gibt eine Zusammenfassung der untersuchten L1 und der entsprechenden Abkürzungen, die im Rahmen der vorliegenden Fallstudie verwendet werden.

²¹² Aus diesem Grund wird dieses Phänomen auch ‚Blocking‘ (Ellis/Wulff 2015: 83) oder ‚proactive inhibition‘ (proaktive Hemmung) genannt (Ellis 2006: 174).

Erstsprachen	Klassifikation im Vergleich zum Deutschen
Englisch (L1E) Italienisch (L1I)	Typologisch ähnlich (L1TA)
Türkisch (L1T) Ungarisch (L1U)	Typologisch verschieden (L1TV)

Tabelle 27: Übersicht zu untersuchten Erstsprachen (L1) der Lernenden.

In der vorliegenden Fallstudie wird der Gebrauch von Präpositionen bzw. PP unabhängig von z. B. ihrer syntaktischen Funktion untersucht. Im Fokus steht die übergeordnete Frage, ob bzw. welche Unterschiede sich im Gebrauch von Präpositionen bei DaF-Lernenden mit typologisch ähnlichen (L1TA) und typologisch verschiedenen L1 (L1TV) zeigen. Es werden folgende Leitfragen angesetzt:

1. (Wie) unterscheidet sich die Korrektheit im Präpositionsgebrauch bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1 (L1TV)?
2. Welche Fehlertypen im Bereich Präpositionsgebrauch zeigen sich bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1(L1TV)? (Wie) unterscheiden sich diese?

Bei beiden Leitfragen wird jeweils zusätzlich das GeR-Gesamtniveau der Lernenden berücksichtigt.

Basierend auf der bisherigen Forschung kann man annehmen, dass die strukturellen Charakteristiken der L1 einen Einfluss auf den Gebrauch von Präpositionen und PP in der L2 Deutsch haben bzw. den Erwerb einer L2 insbesondere zu Beginn steuern. Im L2-Erwerb kommt es bedingt durch die bereits herausgebildete L1 zu Transferphänomenen im Bereich der Sprachverarbeitung. Diese Transferphänomene beziehen sich u.a. auf die im L1-Erwerb aufgebauten Assoziationen zwischen Merkmalen im sprachlichen Input und ihren funktionalen Leistungen (vgl. Mitchell/Myles/Marsden 2013: 106). Im Deutschen fällt darunter z. B. die Assoziation, dass Präpositionen und PP prototypisch Relationen (räumlich, temporal, abstrakt) ausdrücken können. Lernende mit L1 Türkisch (L1T) und L1 Ungarisch (L1U) verfügen aus ihrer L1 nicht über derartige Assoziationen, da diese Sprachen über keine Präpositionen in diesem Sinne verfügen. Lernende mit L1T und L1U stehen folglich im Erwerb der L2 Deutsch vor der Herausforderung, das Konzept der Präpositionen zunächst zu erlernen und diese dann entsprechend ihrer Bedeutung und ihrer Gebrauchskontexte anzuwenden. Dahingegen bestehen bei Lernenden mit L1 Englisch (L1E) und L1 Italienisch (L1I) diese grundlegenden Assoziationen bereits; im Erwerb der L2 Deutsch geht es hier primär darum, aus der L1 bekannte und erworbene Präpositionen und PP sowie deren syntaktische Kontexte (wenn möglich) aufeinander abzubilden und die (teils subtilen) Unterschiede im Gebrauch zu erkennen und anzuwenden (z. B. Rauh 1999). Zum Beispiel verfügt das Deutsche und das Italienische gleichermaßen über Objekt-PP, jedoch liegt nicht immer eine Entsprechung der verwendeten Präpositionen vor (s. grammis²¹³). Fehlende 1:1-Entsprechungen können

²¹³ <https://grammis.ids-mannheim.de/kontrastive-grammatik/3828> (Mai 2019).

Schwierigkeiten in der Wahl zielsprachlich korrekter Präpositionen bedingen (vgl. Blumenthal-Dramé/Kortmann 2013: 293).

Basierend auf den obigen Ausführungen sind Unterschiede im Gebrauch von Präpositionen und PP durch DaF-Lernende mit L1TA und L1TV zu erwarten. Ich gehe grundsätzlich davon aus, dass der Gebrauch von Präpositionen und PP für DaF-Lernende mit L1TV mehr bzw. andere Schwierigkeiten bereitet als DaF-Lernenden mit L1TA. Schwierigkeiten im Präpositionsgebrauch werden durch die Korrektheit und die Fehlertypen im Präpositionsgebrauch operationalisiert. Einen Fehler verstehe ich entsprechend der Definition in Kapitel 3.3.2.2 als eine Abweichung zwischen der Original-Lerneräußerung und der im Korpus verfügbaren Zielhypothese (ZH1). Es werden folgende Hypothesen formuliert, die im Rahmen der korpusgestützten Fallstudie statistisch überprüft werden:

Hypothesen zu Leitfrage 1:

1. Lernende mit L1TA weisen insgesamt eine höhere Korrektheit im Präpositionsgebrauch auf als Lernende mit L1TV.
2. Im Vergleich zu mittleren und höheren Niveaus der Sprachverwendung zeigt sich auf dem Niveau der elementaren Sprachverwendung eine höhere Korrektheit bei Lernenden mit L1TA (im Vergleich zu Lernenden mit L1TV).
3. Unabhängig von der L1 der Lernenden nimmt die Korrektheit im Präpositionsgebrauch mit Anstieg des Kompetenzniveaus zu.

Hypothesen zu Leitfrage 2:

4. Lernende mit L1TV zeigen mehr Schwierigkeiten in der Realisierung von Präpositionen, d. h. sie lassen häufiger Präpositionen aus oder fügen sie in Kontexten hinzu, in denen zielsprachlich keine Präpositionen gefordert sind.
5. Lernende mit L1TA zeigen mehr Schwierigkeiten in der Wahl von Präpositionen.
6. Unabhängig von der L1 der Lernenden nehmen die Schwierigkeiten in der Realisierung von Präpositionen und PP mit Anstieg des Kompetenzniveaus ab.

8.3 Datenauswahl und Datenerhebung

Zur Beantwortung der Leitfragen werden mithilfe der vorhandenen Metadaten im Lernerkorpus MERLIN zwei Subkorpora erstellt, die entsprechend dem Fokus der Fallstudie Daten von Deutschlernenden mit L1TV und L1TA enthalten.

Wie im Kapitel zur Datengrundlage (Kap. 6) beschrieben, enthalten alle Lernertexte im Korpus die Annotationsebene der minimalen Zielhypothese (ZH1) – eine zielsprachliche Rekonstruktion der Lerneräußerungen im Bereich der Orthographie und Grammatik. Dagegen verfügen nicht alle Texte im Korpus über Fehlerannotationen im Bereich der Grammatik (73% der Texte im Korpus sind fehlerannotiert, d. h. sie verfügen über die entsprechende Fehlerannotationsebene EA1). Da in der vorliegenden Fallstudie primär Schwierigkeiten im Gebrauch von Präpositionen und PP fokussiert werden, habe ich mich dafür entschieden, in dieser Fallstudie ausschließlich Lernertexte zu berücksichtigen, die bereits fehlerannotiert sind. D. h. anders als in Fallstudie 2 und Fallstudie 3 wird hier auf eine manuelle (nachträgliche) Annotation der Daten verzichtet. Mit diesem Vorgehen möchte ich exemplarisch aufzeigen,

welche Fragestellungen sich ausschließlich auf Basis bereits vorhandener (Fehler-)Annotationen im Lernerkorpus MERLIN bearbeiten lassen und wie hilfreich Fehlerannotationen in Lernerkorpora für die Bearbeitung spracherwerbstheoretischer Fragestellungen wie der hier vorliegenden sind.

Für die Ermittlung der Lernertexte mit vorliegender EA1-Annotationsebene nutze ich die Funktion der Frequenzanalyse (*Frequency Analysis*) im Korpusrecherchesystem ANNIS (Krause/Zeldes 2016, Zeldes et al. 2009), das als Schnittstelle zum MERLIN-Lernerkorpus fungiert. Mittels der Frequenzanalyse ist es möglich, je L1-Subkorpus die Identifikationsnummern (IDs) der fehlerannotierten Lernertexte zu ermitteln. Basierend darauf²¹⁴ können Abfragen im jeweiligen fehlerannotierten Subkorpus formuliert werden. Tabelle 28 gibt einen Überblick über die auf diese Weise zusammengestellten Subkorpora L1TA und L1TV im Lernerkorpus MERLIN.

L1-Subkorpus	Erstsprache	Anzahl untersch. Lernertexte	Summe Lernertexte	Token	Summe Token
L1TV	Türkisch (L1T)	42	82	4094	13525
	Ungarisch (L1U)	40		9431	
L1TA	Englisch (L1E)	35	51	4663	7353
	Italienisch (L1I)	16		2690	

Tabelle 28: Zusammenstellung der Subkorpora L1TV und L1TA in MERLIN: Überblick.

Da in dieser Fallstudie keine spezifischen Präpositionen bzw. PP im Fokus stehen, sondern der allgemeine Präpositionsgebrauch untersucht wird, wurden in Subkorpora L1TV und L1TA jeweils alle realisierten Präpositionen erhoben. Die Abfrage erfolgte auf Basis der annotierten Wortartenkategorien APPR und APPRART des STTS-Tagsets (Schmid 1994), somit wurden Verschmelzungen und Vollformen gleichermaßen erfasst²¹⁵. Mit dieser Abfrage werden die Gebrauchskontexte von Präpositionen und PP ermittelt, in denen der/die Lernende eine Präposition bzw. PP realisiert hat bzw. eine Einheit als Kategorie APPR bzw. APPRART annotiert wurde, unabhängig davon, ob die Präpositionen bzw. PP im entsprechenden Kontext zielsprachlich gefordert und/oder korrekt ist. Weiterhin werden mit einer zusätzlichen Abfrage alle realisierten, aber fehlerannotierten Präpositionen innerhalb der L1-Subkorpora

²¹⁴ Die so ermittelten IDs je Subkorpus wurden gespeichert und im Anschluss für weitere Korpusabfragen eingesetzt.

²¹⁵ Die Kategorie APPRART entspricht laut den Richtlinien einer Präposition mit Artikel (*Er geht zum Arzt*) (Schiller/Teufel/Stöckert 1999: 64). Die Kategorie APPR entspricht laut STTS-Richtlinien (Schiller/Teufel/Stöckert 1999: 64) einer vorangestellten Präposition in ihrer nicht verschmolzenen Form (z. B. *auf dem Tisch*), dem linken Bestandteil einer Zirkumposition (*um Himmels willen*), präpositionsähnlichen Adjektiven (*gleich seinem Bruder*) bzw. präpositionsähnlichen Partizipien (*ausgenommen dreier Burgen*). Eine manuelle Durchsicht der erhobenen Korpusdaten ergab, dass im Datensatz der vorliegenden Fallstudie weder präpositionsähnliche Adjektive noch präpositionsähnliche Partizipien oder die Zirkumpositionen (*um [...] willen*) vorliegen, die nach der GDS (1997: 2085) die „einzige eindeutige“ Zirkumposition im Deutschen ist. Somit kann angenommen werden, dass im erhobenen Datensatz lediglich vorangestellten Präposition und deren Verschmelzungen vorliegen.

erhoben. Für diese Abfrage werden die Ebenen der Wortartenannotation und der Fehlerannotation kombiniert und die entsprechenden Fehlerannotationskategorien abgefragt (s. Wisniewski et al. 2014: 12). Um PP-Kontexte zu berücksichtigen, in denen Lernende zielsprachlich geforderte Präpositionen ausgelassen haben, werden zusätzlich alle Instanzen des entsprechenden Fehlertyps (Auslassung) je L1-Subkorpus ermittelt. Diese Instanzen werden mithilfe separater Abfragen ermittelt, da im Falle von Auslassungen keine Wortannotationen von Präpositionen vorliegen, und somit nicht auf entsprechende Annotationen in den Korpusabfragen rekuriert werden kann. Die Berücksichtigung von Auslassungen ist äußerst wichtig, um die Grundgesamtheit aller zielsprachigen und nicht-zielsprachigen PP-Kontexte zu ermitteln, die in den weiterführenden Analysen berücksichtigt werden. Die Korpusabfragen sind in Tabelle 29 notiert, Tabelle 30 gibt einen Überblick über die Ergebnisse der entsprechenden Abfragen je L1-Subkorpus. Der Begriff *PP-Kontexte* schließt im Folgenden sowohl den zielsprachlichen als auch den nicht-zielsprachlichen Gebrauch von Präpositionen bzw. PP ein und inkludiert auch Kontexte, in denen keine Präposition bzw. PP realisiert wurde, zielsprachlich jedoch eine gefordert war.

	Korpusabfrage	Beschreibung
1	<code>meta:__author_id=/(.....)²¹⁶/ & tok_pos=/APPR APPRART/</code>	Ermittlung aller Einheiten, die als Präpositionen (Vollform o. Verschmelzung) annotiert wurden = Realisierte Präpositionen
2	<code>meta:__author_id=/(.....)/ & G_Prep_type=/ch pos ad²¹⁷/ _i_tok_pos=/APPR APPRART/</code>	Ermittlung aller Einheiten, bei denen eine Fehlerannotation im Bereich Präpositionsgebrauch vorliegt und die gleichzeitig eine annotierte Präposition (Vollform o. Verschmelzung) repräsentieren = Fehlerannotierte (realisierte) Präpositionen
3	<code>meta:__author_id=/(.....)/ & G_Prep_type="o"</code>	Ermittlung aller Instanzen des Fehlertyps Auslassung im Präpositionsgebrauch = Fehlerannotierte ausgelassene Präpositionen

Tabelle 29: Korpusabfragen zur Fallstudie Erstsprache der Lernenden.

²¹⁶ Zwischen den Schrägstrichen wurden alle IDs der fehlerannotierten Lernertexte je L1-Subkorpus eingefügt. Ich verzichte an dieser Stelle auf eine Auflistung dieser IDs.

²¹⁷ Zu den Fehlerannotationskategorien s. Kapitel 6 und Kapitel 8.5. Die Fehlerannotationskategorien *merge* und *split* sind nicht in den entsprechenden L1-Subkorpora enthalten und werden aus diesem Grund hier nicht aufgeführt.

L1-Subkorpus	Erstsprache	1	2	3
		Realisierte Präpositionen	Fehlerannotierte (realisierte) Präpositionen ²¹⁸	Fehlerannotierte ausgelassene Präpositionen
L1TV	Türkisch (L1T)	246	15	20
	Ungarisch (L1U)	627	27	19
L1TA	Englisch (L1E)	375	29	11
	Italienisch (L1I)	188	10	2

Tabelle 30: Ergebnis der Korpusabfragen je L1-Subkorpus (nicht bereinigt).

Aus dem obigen Ergebnis der Korpusabfragen ergibt sich eine Grundgesamtheit von 1488²¹⁹ PP-Kontexten, in der sowohl zielsprachlich geforderte als auch zielsprachlich nicht geforderte PP-Kontexte enthalten sind. Tabelle 31 gibt einen Überblick über die Grundgesamt der ermittelten PP-Kontexte je L1-Subkorpus.

L1-Subkorpus	Erstsprache	PP-Kontexte		Summe
L1TV	Türkisch (L1T)	266	912	1488
	Ungarisch (L1U)	646		
L1TA	Englisch (L1E)	386	576	
	Italienisch (L1I)	190		

Tabelle 31: PP-Kontexte je L1-Subkorpus (nicht bereinigt).

Die auf Basis der Korpusabfragen ermittelten PP-Kontexte wurden jeweils auf zwei Arten exportiert:

1) Für den Export der PP-Kontexte wurde der im Korpusrecherchesystem ANNIS integrierte GridExporter verwendet. Dabei wurden folgende Daten exportiert: die Original-Lerneräußerung, die entsprechende ZH1, die Identifikationsnummer, die Erstsprache, das GeR-Gesamtniveau sowie, wenn vorliegend, die Fehlerannotation im Bereich Präpositionsgebrauch.

2) Die ermittelten Korpustreffer der jeweiligen Korpusabfragen wurden einer Frequenzanalyse unterzogen, die ebenfalls im Korpusrecherchesystem ANNIS integriert ist. Mit der Frequenzanalyse ist es möglich, für die Korpustreffer Häufigkeitsverteilungen der mit der Korpusabfrage ermittelten Kategorien (z. B. APPR/APPRART oder Fehlerannotationskategorien) unter Berücksichtigung ausgewählter Metadaten (z. B. L1 oder GeR-Gesamtniveau) in einer Tabelle ausgeben zu lassen. Der Output der Frequenzanalyse wurde exportiert.

Die exportierten Daten aus 1) werden verwendet, um den Präpositionsgebrauch im Kontext zu betrachten sowie Verwendungsbeispiele je L1-Subkorpus zu extrahieren. Die Ergebnisse der Frequenzanalyse 2) werden für die Datenauswertung in R verwendet.

²¹⁸ Die fehlerannotierten (realisierten) Präpositionen sind in den Angaben realisierten Präpositionen (1) enthalten, und sind hier nochmal separat aufgeführt.

²¹⁹ Die Summe aus realisierten Präpositionen und fehlerannotierten, ausgelassenen Präpositionen je L1-Subkorpus.

8.4 Datenaufbereitung

Vor der Datenauswertung wurden die Exporte aller fehlerannotierten PP-Kontexte gesichtet und die Zuordnung der vorliegenden Fehlerannotationskategorien überprüft, wobei ich mich an den entsprechenden Vergaberichtlinien von MERLIN orientiert habe (s. Wisniewski et al. 2014). In einigen Fällen waren Korrekturen vorliegender Fehlerannotationen notwendig. Diese PP-Kontexte bzw. die ihnen zugeordneten Fehlerannotationskategorien wurden in der Phase der Datenaufbereitung manuell bereinigt, d. h. fehlerhafte Fehlerannotationen²²⁰ wurden identifiziert und für die weitere Analyse korrigiert. Für die Datenauswertung wurden lediglich die korrigierten Fehlerannotationen im Präpositionsgebrauch ausgewertet²²¹.

Ein Lernertext (ID: 1061_0120358) musste von der weiteren Analyse ausgeschlossen werden, da im Korpus ein Problem mit der Alinierung von Original-Lerneräußerung und ZH1 vorlag. Somit wurden alle in diesem Lernertext ermittelten PP-Kontexte aus dem Datensatz entfernt. Die folgende Tabelle (Tab. 32) gibt einen Überblick über die bereinigten Daten, die in der weiterführenden Datenauswertung berücksichtigt werden.

L1-Subkorpus	Erstsprache	Realisierte Präpositionen (davon fehlerannotierte Präpositionen)		Ausgelassene Präpositionen		Summe ²²²	
L1TV	Türkisch (L1T)	246 (22)	873 (73)	20	37	266	910
	Ungarisch (L1U)	627 (51)		17		644	
L1TA	Englisch (L1E)	368 (40)	556 (50)	8	10	376	566
	Italienisch (L1I)	188 (10)		2		190	

Tabelle 32: PP-Kontexte je L1-Subkorpus (bereinigt).

Es erfolgte keine weitere Aufbereitung der bereinigten Korpusdaten. Die im Lernerkorpus ermittelten PP-Kontexte werden entsprechend den oben angeführten Leitfragen einer quantitativen Analyse unterzogen, die im Folgenden ausführlich beschrieben wird. Insgesamt werden 1476 PP-Kontexte (L1TV und L1TA) ausgewertet.

8.5 Datenauswertung

Im Rahmen der Datenauswertung wurde ermittelt, wie viele unterschiedliche Lernertexte, d. h. Lernende je L1-Subkorpus durch die erhobenen PP-Kontexte repräsentiert sind. Die Ergebnisse sind in der nachfolgenden Tabelle (Tab. 33) abgebildet. Der Tabelle kann entnommen werden, dass je L1-Subkorpus eine verschiedene Anzahl an Lernenden vorliegt, insgesamt sind 131 Lernende (L1TV und L1TA) repräsentiert. Somit liegen im Datensatz keine

²²⁰ Zu den fehlerhaften Annotationen zählen hier inkorrekte Fehlerannotationskategorien und Fehlerannotationen trotz nicht vorliegender Fehler.

²²¹ An dieser Stelle möchte ich erwähnen, dass je L1-Subkorpus lediglich die Fehlerannotationen im Präpositionsgebrauch manuell bereinigt wurden. Es kann somit nicht vollständig ausgeschlossen werden, dass in den nicht-fehlerannotierten PP-Kontexten vereinzelt doch Fehler im Präpositionsgebrauch vorliegen.

²²² Die Summe ergibt sich aus den Angaben zu realisierten Präpositionen und den Angaben zu ausgelassenen Präpositionen.

Datenausreißer vor, d. h. die ermittelten PP-Kontexte gehen nicht nur auf wenige einzelne Lernende zurück.

L1-Subkorpus	Erstsprache	PP-Kontexte gesamt	Summe PP- Kontexte	AL	Summe AL
L1TV	Türkisch (L1T)	266	910	41	81
	Ungarisch (L1U)	644		40	
L1TA	Englisch (L1E)	376	566	34	50
	Italienisch (L1I)	190		16	

Tabelle 33: Verteilung der PP-Kontexte auf unterschiedliche Lernertexte je L1-Subkorpus. AL=Anzahl Lernertexte.

Die ermittelten PP-Kontexte wurden entsprechend den obig formulierten Leitfragen ausgewertet. An dieser Stelle seien die Leitfragen nochmals aufgeführt:

1. (Wie) unterscheidet sich die Korrektheit im Präpositionsgebrauch bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1 (L1TV)?
2. Welche Fehlertypen im Bereich Präpositionsgebrauch zeigen sich bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1 (L1TV)? (Wie) unterscheiden sich diese?

Für die Beantwortung dieser Leitfragen wurde je L1-Subkorpus jeweils zusätzlich das GeR-Gesamtniveau der Lernenden berücksichtigt.

8.5.1 Analysemethoden

Die ermittelten PP-Kontexte beider L1-Subkorpora werden mittels der computerunterstützten Fehleranalyse (computer-aided error analysis, CEA) (Dagneaux/Denness/Granger 1998) sowie der kontrastiven Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (Granger 1996, 2015) ausgewertet. Die nachfolgende Tabelle (Tab. 34) gibt einen Überblick über die in der vorliegenden Fallstudie angewandten Analysemethoden und die damit verbundenen Ziele. Nähere Informationen zur Anwendung dieser Methoden im Kontext der einzelnen Fallstudien bietet Kapitel 7.1.

Analysemethode	Ziel
Computerunterstützte Fehleranalyse	Ermittlung der Abweichungen (bzw. Korrektheit) im Präpositionsgebrauch (in zielsprachlich geforderten und zielsprachlich nicht geforderten PP-Kontexten) je L1-Subkorpus (L1TV und L1TA) und GeR-Gesamtniveau.
Kontrastive Interlanguage-Analyse	Kontrastiver Vergleich der Ergebnisse der computerunterstützten Fehleranalyse je GeR-Gesamtniveau und L1-Subkorpus (L1TV und L1TA).

Tabelle 34: Analysemethoden der Fallstudie zur Variable *Erstsprache der Lernenden*: Überblick.

Im folgenden Kapitel werden die Ergebnisse der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse und ihre statistischen Auswertungen präsentiert.

8.5.2 Ergebnisse

In diesem Kapitel werden die Ergebnisse der quantitativen Datenauswertung vorgestellt, die eingangs formulierten Hypothesen überprüft und die entsprechenden Leitfragen beantwortet. Das Kapitel ist entsprechend der eingangs formulierten Leitfragen gegliedert: Der Analyse und dem Vergleich der Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Leitfrage 1) folgt die Analyse und der Vergleich der Fehlertypen im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Leitfrage 2). Dabei wird jeweils überprüft, ob bzw. inwieweit die L1 der Lernenden einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch ausübt.

8.5.2.1 Auswertung Korrektheit im Präpositionsgebrauch

Die quantitative Auswertung der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse liefern in Bezug auf die Korrektheit im Präpositionsgebrauch drei zentrale Ergebnisse:

- 1) Insgesamt kann eine hohe Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch ermittelt werden.
- 2) Ein signifikanter Effekt der L1 auf die Korrektheit im Präpositionsgebrauch kann nicht ermittelt werden; ein deutlicher Unterschied zwischen L1TV und L1TA ist lediglich im fortgeschrittenen GeR-Gesamtniveau B2 zu beobachten.
- 3) Es lässt sich ein signifikanter Effekt des GeR-Gesamtniveaus auf die Korrektheit im Präpositionsgebrauch ermitteln: Mit steigendem GeR-Gesamtniveau nimmt die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch signifikant zu, und das unabhängig von der L1 der Lernenden.

Die folgende Tabelle (Tab. 35) gibt einen Überblick über die Ergebnisse (absolut und relativ) zur Korrektheit im Präpositionsgebrauch je L1-Subkorpus.

L1-Subkorpus	Erstsprache	PP-Kontexte (Gesamt)		Inkorrekte PP-Kontexte			Korrekte PP-Kontexte		
L1TV	Türkisch (L1T)	266	100%	42	16%	110	224	84%	800
	Ungarisch (L1U)	644	100%	68	11%	(12%)	576	89%	(88%)
L1TA	Englisch (L1E)	376	100%	48	13%	60	328	87%	506
	Italienisch (L1I)	190	100%	12	6%	(11%)	178	94%	(89%)

Tabelle 35: Korrektheit im Präpositionsgebrauch je L1-Subkorpus: Überblick.

Die Zusammenfassung in Tabelle 35 zeigt, dass die Lernenden unabhängig von ihrer L1 insgesamt einen hohen Anteil an korrekt gebrauchten Präpositionen bzw. PP-Kontexten zeigen (88% bei L1TV bzw. 89% bei L1TA). Insgesamt können 170 inkorrekte Verwendungen von Präpositionen ermittelt werden, das entspricht einem Anteil von 11,52% im gesamten Datensatz der vorliegenden Fallstudie. Beispiele für den inkorrekten Präpositionsgebrauch je L1-Subkorpus sind im Folgenden aufgeführt:

L1TA:

1. *Ich freue mich um Ihrer Angebot* (ID: 1023_0101841)
ZH1: Ich freue mich über Ihr Angebot
2. *Jetzt ich suche für eine Wohnung* (ID: 1071_0024702)
ZH1: Jetzt suche ich eine Wohnung
3. *ich bin für diese Stelle sehr interessiert.* (ID: 1023_0109717)
ZH1: ich bin an dieser Stelle sehr interessiert.
4. *Nach dem Abendessen sind wir in Zentralplatz [...] gegangen* (ID: 1061_0120882)
ZH1: Nach dem Abendessen sind wir auf den Zentralplatz [...] gegangen

L1TV:

1. *im Ausserhalb des Stadt gibts nicht nur schone Wohnungen* (ID: 1031_0003414)
ZH1: Außerhalb der Stadt gibts nicht nur schöne Wohnungen
2. *Ich warte Dich* (ID: 1071_0243591)
ZH1: Ich warte auf Dich
3. *Aber man muß auch an der Kultur des Gastlandes kennenzulernen.* (ID: 1031_0003231)
ZH1: Aber man muss auch die Kultur des Gastlandes kennenlernen.
4. *Nechsten Monat möchte ich euch gehen.* (ID: 1071_0024873)
ZH1: Nächsten Monat möchte ich zu euch gehen.

In der folgenden Tabelle (Tab. 36) ist die Korrektheit im Präpositionsgebrauch je L1-Subkorpus und getrennt nach GeR-Gesamtniveau dargestellt, so wie im Datensatz vorliegend. Zur besseren Übersicht sind in Tabelle 36 die L1-Subkorpora L1TV/L1TA hier nur zusammenfassend angegeben.

L1-Subkorpus	PP-Kontexte	GeR-Gesamtniveau								
		A1	A2	A2+	B1	B1+	B2	B2+	C1	Summe
L1TV	Gesamt	5	88	20	54	180	376	88	99	910
	inkorrekt	3	24	2	6	18	43	11	3	110
	korrekt	2	64	18	48	162	333	77	96	800
L1TA	Gesamt	9	53	38	70	103	277	16	0	566
	inkorrekt	2	13	4	13	11	17	0	0	60
	korrekt	7	40	34	57	92	260	16	0	506

Tabelle 36: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau.

Die Werte in Tabelle 36 zeigen, dass die einzelnen GeR-Gesamtniveaus nicht gleichermaßen im Datensatz repräsentiert sind. Die mittleren bis höheren Kompetenzniveaus B1+ und B2 sind besonders stark vertreten, das GeR-Gesamtniveau A1 dahingegen nahezu gar nicht und das GeR-Gesamtniveau C1 nur im Subkorpus L1TV. Zu beachten ist allerdings, dass je L1-Subkorpus jeweils zwei L1 abgedeckt sind, die ihrerseits ebenfalls unterschiedliche Distributionen von PP-Kontexten je GeR-Gesamtniveau aufweisen. Die folgende Tabelle (Tab. 37) gibt hierzu einen detaillierten Überblick, wobei Kreuze diejenigen GeR-Niveaus markieren, in denen PP-

Kontexte im Datensatz vorliegen. Zur besseren Übersicht wird an dieser Stelle auf die Angabe der Anzahl von PP-Kontexten je L1 verzichtet.

L1-Subkorpus	Erstsprache	GeR-Gesamtniveau							
		A1	A2	A2+	B1	B1+	B2	B2+	C1
L1TV	Türkisch (L1T)	x	x	x	x	x	x		
	Ungarisch (L1U)		x		x	x	x	x	
L1TA	Englisch (L1E)	x	x	x	x	x	x		
	Italienisch (L1I)		x		x	x	x	x	

Tabelle 37: PP-Kontexte je Erstsprache und GeR-Gesamtniveau: Überblick.

Wie man Tabelle 37 entnehmen kann, sind die GeR-Gesamtniveaus A2, B1, B1+ und B2 in jedem L1-Subkorpus vertreten. Diese GeR-Gesamtniveaus würden sich folglich für eine vergleichende Analyse der L1-Subkorpora anbieten. Aus der vorherigen Tabelle (Tab. 36) geht jedoch hervor, dass im L1-Subkorpus L1TV im Niveau B1 vergleichsweise nur sehr wenige inkorrekte PP-Kontexte im Datensatz enthalten sind. Mit Blick auf die weitere Auswertung der Fehlertypen stellt dies ein Problem dar, da hier nur sehr einschränkende Aussagen möglich wären. Für die weitere Auswertung habe ich mich daher dafür entschieden, lediglich die GeR-Gesamtniveaus A2, B1+ und B2 zu berücksichtigen²²³.

Die folgende Tabelle (Tab. 38) gibt eine Übersicht über die Anzahl korrekter und inkorrekt PP-Kontexte je L1-Subkorpus und ausgewähltem GeR-Gesamtniveau. Die Anzahl unterschiedlicher Lernertexte, d. h. Lernender je GeR-Gesamtniveau und L1-Subkorpus ist ebenfalls in der Tabelle angegeben.

L1-Subkorpus		GeR-Gesamtniveau						
		A2		B1+		B2		Summe
		PP-Kontexte	AL	PP-Kontexte	AL	PP-Kontexte	AL	PP-Kontexte
L1TV	inkorrekt	24 (27%)	21	18 (10%)	11	43 (14%)	23	85 (13%)
	korrekt	64 (73%)		162 (90%)		333 (86%)		559 (87%)
L1TA	inkorrekt	13 (25%)	11	11 (11%)	6	17 (6%)	16	41 (9%)
	korrekt	40 (75%)		92 (89%)		260 (94%)		392 (91%)

Tabelle 38: Korrekte und inkorrekte PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau: Überblick. AL=Anzahl unterschiedlicher Lernertexte.

Insgesamt werden je L1-Subkorpus 644 (L1TV) bzw. 433 (L1TA) PP-Kontexte ausgewertet. Im Datensatz sind insgesamt 55 Lernende (L1TV) bzw. 33 Lernende (L1TA) repräsentiert. Die Anzahl der PP-Kontexte sowie die Anzahl unterschiedlicher Lernertexte (und damit

²²³ Eine Durchsicht der inkorrekten PP-Kontexte auf dem Niveau B1 zeigt, dass einige Fehlertypen lediglich ein einziges Mal im Datensatz enthalten sind. In den Niveaus A2, B1+ und B2 sind Fehlertypen dahingehen mehrfach vertreten. Um die GeR-Gesamtniveaus bei der Analyse zur Korrektheit und zu den Fehlertypen konstant zu halten, habe ich mich dazu entschlossen, das Niveau B1 ebenfalls für die Auswertung zur Korrektheit im Präpositionsgebrauch auszuschließen.

unterschiedlicher Lernender) innerhalb der ausgewählten GeR-Gesamtniveaus ist damit sehr eingeschränkt. Dennoch kann der Vergleich des Präpositionsgebrauchs in den vorliegenden L1-Subkorpora innerhalb der ausgewählten GeR-Gesamtniveaus relevante Einblicke gewähren und Tendenzen im Verlauf des Kompetenzanstiegs aufzeigen. Weiterführende Untersuchungen in Lernerkorpora können an diese Ergebnisse anknüpfen.

Der vorliegende Datensatz wird statistisch ausgewertet. Hierfür wähle ich die Regressionsanalyse und berechne gemischte Modelle (Näheres zur statistischen Auswertung s. Kap. 7.2). Als Zielvariable (abhängige Variable) wähle ich die Korrektheit im Präpositionsgebrauch. Als feste Faktoren (unabhängige kategoriale Variablen) wähle ich das L1-Subkorpus (L1.Zf²²⁴) sowie das GeR-Gesamtniveau (Kompetenzniveau). Der/Die einzelne Lernende wird jeweils als Zufallsfaktor (random effect) festgelegt (1|Lerner). Auf diese Weise kann der individuelle Einfluss des/der Lernenden herausgerechnet werden. Im ersten Regressionsmodell integriere ich eine Interaktion zwischen den beiden kategorialen Variablen (GeR-Gesamtniveau und L1), um zu überprüfen, ob sich diese Variablen gegenseitig beeinflussen. Das entsprechende Regressionsmodell (RM1) und sein Output sind im Folgenden angegeben:

RM1:

```
glmer(Korrekteit ~ L1.Zf*Kompetenzniveau + (1|Lerner), data = dat, family = "binomial",
control = glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.1628	0.3541	3.283	0.001025
L1.ZfL1TV	-0.1495	0.4381	-0.341	0.733014
KompetenzniveauB1+	1.0461	0.5089	2.056	0.039824
KompetenzniveauB2	1.6198	0.4453	3.638	0.000275
L1.ZfL1TV:KompetenzniveauB1+	0.2154	0.6341	0.340	0.734092
L1.ZfL1TV:KompetenzniveauB2	-0.5181	0.5471	-0.947	0.343644

Die Ergebnisse der statistischen Auswertungen zeigen, dass zwischen L1-Subkorpus und dem GeR-Gesamtniveau keine allgemeine Interaktion vorliegt ($p=0.3732$), d. h. die L1 und das GeR-Gesamtniveau der Lernenden beeinflussen sich nicht gegenseitig in ihrem Einfluss auf die Korrektheit im Präpositionsgebrauch. Die Werte des Modell-Outputs (s.o.) lassen den Schluss zu, dass die Erstsprache keinen Effekt ausübt, jedoch ein Effekt des GeR-Gesamtniveaus vorzuliegen scheint. Das Modell (RM1) ist in Abbildung 32 visualisiert. Auf der Y-Achse ist die Wahrscheinlichkeit²²⁵ für Korrektheit im Präpositionsgebrauch abgebildet, auf der x-Achse finden sich die einzelnen GeR-Gesamtniveaus; Lernende mit L1TV sind als gestrichelte Linie repräsentiert, Lernende mit L1TA als durchgezogene Linie. Für jedes GeR-Gesamtniveau ist

²²⁴Die Abkürzung *L1.Zf* steht für die Zusammenfassung der einzelnen L1 zu den L1-Subkorpora L1TV und L1TA.

²²⁵ Beispiel: Der Wert 0.75 repräsentiert 75%.

der Mittelwert zur Wahrscheinlichkeit der Korrektheit angegeben (Kreis) sowie die Standardabweichung zum jeweiligen Mittelwert (senkrechte Linie mit waagerechten Linien).

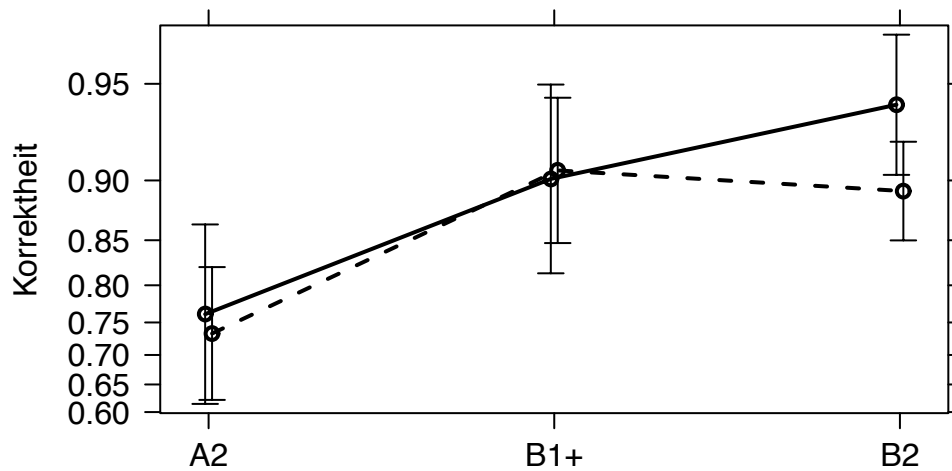


Abbildung 32: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM1. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.

Abbildung 32 illustriert, dass die Korrektheit im Präpositionsgebrauch (sowie im Datensatz vorliegend) insgesamt sehr hoch ist. Inwiefern beim Gebrauch von Präpositionen Vermeidungsstrategien eine Rolle spielen, kann hier nicht beurteilt werden. Es wäre aber denkbar, dass Lernende insgesamt vorwiegend diejenigen Präpositionen verwenden, bei deren Gebrauch sie sich sicher fühlen und Präpositionen vermeiden, deren Gebrauch sie noch nicht beherrschen. Dies würde den hohen Anteil an Korrektheit im Präpositionsgebrauch ebenfalls erklären. Die typologische Nähe der L1 zu L2 könnte bei Vermeidungsstrategien eine Rolle spielen²²⁶.

In Bezug auf den Unterschied zwischen den L1-Gruppen zeigt Abbildung 32, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch auf den Niveaus A2 und B1+ für Lernende mit L1TV und L1TA sehr ähnlich ist. Hier zeigen sich fast keine Unterschiede, die Markierungen der Standardabweichungen überlappen jeweils die entsprechenden Mittelwerte; auf dem GeR-Gesamtniveau B1+ überlappen sich gar die Mittelwerte beider L1-Gruppen. Auf dem GeR-Gesamtniveau B2 dahingegen lässt Abbildung 32 erkennen, dass die Wahrscheinlichkeit für Korrektheit bei Lernenden mit LTA höher liegt als bei Lernenden mit L1TV, hier überlappen die Markierungen der Standardabweichungen (knapp) nicht die entsprechenden Mittelwerte. Lernende mit L1TA zeigen auf der höheren Stufe selbstständiger Sprachverwendung (Niveau B2) (Europarat 2001: 40f.) demnach eine leicht größere Unsicherheit im Gebrauch von deutschen Präpositionen und PP im Vergleich zu Lernenden mit L1TA. Weiterhin lässt sich der obigen Abbildung entnehmen, dass die Wahrscheinlichkeit für Korrektheit bei Lernenden mit L1TV zwischen den GeR-Gesamtniveaus B1+ und B2 nahezu

²²⁶ Zu Vermeidungsstrategien beim Gebrauch von Präpositionen und PP erwähnt Lütke (2011: 107) in einem Forschungsüberblick, dass kindliche Lernende mit L1 Türkisch zur Darstellung räumlicher Beziehungen präpositionale Wendungen meiden und stattdessen Verb-Adverb-Verbindungen wählen. Inwiefern diese Strategie auch bei den im hier vorliegenden Datensatz repräsentierten Lernenden greift, kann nicht ermittelt werden, da ausschließlich Präpositionen bzw. PP erhoben wurden.

stagniert bzw. etwas abnimmt während sie bei Lernenden mit L1TA weiter ansteigt. Ein Unterschied in der Korrektheit im Präpositionsgebrauch zwischen Lernenden mit unterschiedlichen L1 zeigt sich somit lediglich auf dem höheren Kompetenzniveau B2. Vermutlich lässt sich hier die Tendenz beobachten, dass Lernende mit L1TV im Spracherwerbsprozess eine längere Zeit benötigen als Lernende mit L1TA, um deutsche Präpositionen und PP zu erwerben und diese zielsprachlich angemessen zu produzieren. Ähnliche Tendenzen im unterschiedlichen Erwerbtempo bei der Realisierung von (lokalen und direktionalen) Präpositionen beobachtet auch Turgay (2010a: 278) bei Kindern mit L1 Türkisch und L1 Italienisch²²⁷.

Wie oben erwähnt, wurde im Modell RM1 keine Interaktion zwischen L1 und GeR-Gesamtniveau ermittelt. Aus diesem Grund wird die Interaktion aus dem Modell entfernt. In einem modifizierten Modell (RM1a) werden lediglich die Einzeleffekte des GeR-Gesamtniveaus auf die Korrektheit im Präpositionsgebrauch berechnet (RM1a):

RM1a:

```
glmer(Korrekttheit ~ L1.Zf + Kompetenzniveau + (1|Lerner), data = dat, family = "binomial",
control = glmerControl(optimizer = "bobyqa"))
```

Der Modell-Output ist im Folgenden angegeben, die Visualisierung des Modell-Outputs (RM1a) ist unten (Abb. 33) abgebildet.

RM1a:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.3053	0.2647	4.931	8.18e-07
L1.ZfL1TV	-0.3555	0.2311	-1.538	0.124056
KompetenzniveauB1+	1.1964	0.3135	3.816	0.000135
KompetenzniveauB2	1.2843	0.2656	4.835	1.33e-06

²²⁷ Wegener (1995b: 7) kann ähnliches für den Erwerb des deutschen Genus bei Kindern mit L1 Türkisch und Russisch bzw. Polnisch feststellen: Kinder mit L1 Türkisch – einer genuslosen Sprache – zeigen hier im Vergleich zu Kindern mit L1 Russisch/Polnisch – Sprachen mit drei Genera und in dieser Hinsicht ähnlich zum Deutschen – ein langsames Erwerbtempo.

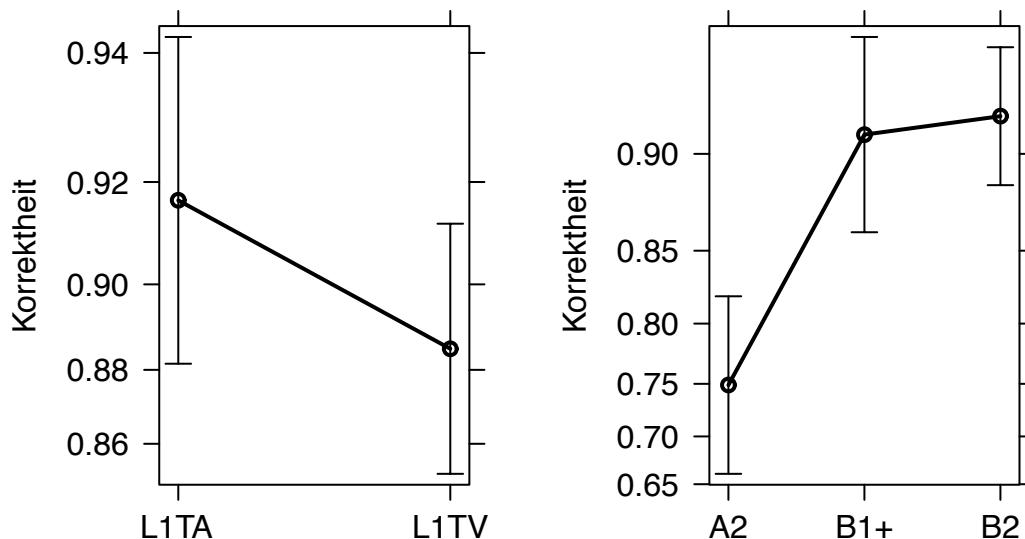


Abbildung 33: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM1a.

Zunächst lässt sich dem Modell-Output und Abbildung 33 entnehmen, dass kein Effekt des L1-Subkorpus auf die Korrektheit im Präpositionsgebrauch ermittelt werden kann. Das bedeutet, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch unabhängig vom L1-Subkorpus ähnlich hoch ist, auch wenn sich tendenziell eine höhere Wahrscheinlichkeit bei Lernenden mit L1TA zeigt (s. Abb. 33). Weiterhin illustriert Abbildung 33, dass mit Zunahme des GeR-Gesamtniveaus der Lernenden auch die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch ansteigt. Die Berechnungen deuten darauf hin, dass dies unabhängig von der L1 der Lernenden zu beobachten ist. Ein Anstieg in der Korrektheit ist zu erwarten, da mit steigender Sprachkompetenz und damit verbundenem steigendem Sprachwissen i.d.R. insgesamt auch eine Reduktion sprachlicher Fehler einhergeht²²⁸.

Aus Abbildung 33 wird zudem ersichtlich, dass die allgemeine Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch zwischen dem GeR-Gesamtniveau A2 und B1+ stärker ansteigt als zwischen den GeR-Gesamtniveaus B1+ und B2. Dies lässt sich vermutlich u.a. dadurch erklären, dass die sprachlichen Kompetenzen, die durch die GeR-Gesamtniveaus B1+ und B2 repräsentiert sind, näher beieinander liegen, d. h. sich in ihren sprachlichen Ausprägungen stärker ähneln als die GeR-Gesamtniveaus A2 und B1+ (Europarat 2001: 41). Es ist somit plausibel anzunehmen, dass sich diese ‚Nähe‘ zwischen B1+ und B2 auch bei der Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch zeigt. Da im Datensatz die GeR-Gesamtniveaus A2+ und B1 nicht repräsentiert sind, lassen sich an dieser Stelle leider keine Aussagen zur Wahrscheinlichkeit der Korrektheit im Präpositionsgebrauch in diesen GeR-Gesamtniveaus treffen. Untersuchungen zu weiteren Kompetenzniveaus auf Grundlage einer umfangreicheren Datenbasis könnten hier weitere Erkenntnisse liefern. Die Ergebnisse der vorliegenden quantitativen Auswertung zeigen, dass die Korrektheit im Präpositionsgebrauch

²²⁸ Zum Phänomen des *U-shaped learning* (bei dem sprachliche Fehler mit steigender Kompetenz aufgrund von Übergeneralisierungen zunächst zunehmen bevor sie wieder abnehmen) aus Perspektive der kognitiven Spracherwerbtheorien siehe z. B. Elsen (2014: 215ff.) und Ellis (2003: 87f.).

signifikant durch das GeR-Gesamtniveau der Lernenden beeinflusst wird, nicht jedoch durch die L1 der Lernenden.

Für Evaluationszwecke wurde das Regressionsmodell entsprechend modifiziert (RM1b, s.u.) und lediglich der Einzeleffekt des GeR-Gesamtniveaus auf die Korrektheit im Präpositionsgebrauch berücksichtigt. Das modifizierte Modell ist mit seinem Output im Folgenden angegeben:

RM1b:

```
glmer(Korrektheit ~ Kompetenzniveau + (1|Lerner), data = dat, family = "binomial", control =  
glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.0814	0.2194	4.930	8.24e-07
KompetenzniveauB1+	1.2030	0.3204	3.755	0.000173
KompetenzniveauB2	1.3017	0.2705	4.812	1.50e-06

Wie die Werte im Modell-Output illustrieren, lässt sich ein signifikanter Anstieg der Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch mit steigendem GeR-Gesamtniveau ermitteln. Weitere statistische Berechnungen zeigen, dass auch ein allgemeiner signifikanter Effekt des GeR-Gesamtniveaus vorliegt ($p= 4.115e-06$). Die Evaluation des modifizierten Modells RM1b zeigt jedoch, dass dieses Modell (d. h. der Effekt des GeR-Gesamtniveaus unter Berücksichtigung des/der einzelnen Lernenden als Zufallsfaktor) die Varianz in der Korrektheit im Präpositionsgebrauch lediglich zu 4,3% erklären kann ($R^2 = 0.04263937$). Die Werte zeigen insgesamt also ein sehr niedriges Erklärungspotenzial des Modells bzw. ein sehr niedriges Erklärungspotenzial des Einzeleffekts der sprachlichen Kompetenz auf die Korrektheit im Präpositionsgebrauch in den analysierten Daten. Ein signifikanter Effekt des GeR-Gesamtniveaus lässt sich zwar nachweisen, der R^2 -Wert deutet jedoch darauf hin – und dies ist plausibel –, dass die Korrektheit im Präpositionsgebrauch nicht ausschließlich von dieser Variable beeinflusst wird.

Zusammenfassung

Die erste Leitfrage kann anhand der Auswertungen wie folgt beantwortet werden:

1. (Wie) unterscheidet sich die Korrektheit im Präpositionsgebrauch bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1 (L1TV)?

Für beide L1-Subkorpora zeigt sich insgesamt eine hohe Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch, die mit Anstieg des GeR-Gesamtniveaus weiter zunimmt. Für die Auswertungen wurden die GeR-Gesamtniveaus A2, B1+ und B2 berücksichtigt. Den statistischen Auswertungen nach hat die L1 der Lernenden keinen Effekt auf die Korrektheit im Präpositionsgebrauch. Dahingegen lässt sich ein signifikanter Effekt des GeR-Gesamtniveaus auf die Korrektheit im Präpositionsgebrauch feststellen. Somit kann auf Basis der Auswertungen des vorliegenden Datensatzes geschlossen werden, dass sich die

Korrektheit im Präpositionsgebrauch durch das GeR-Gesamtniveau der Lernenden besser vorhersagen lässt als durch die typologische Nähe der jeweiligen L1 zur L2 Deutsch. Die Evaluation der entsprechenden statistischen Modelle lässt jedoch insgesamt eher ein niedriges Erklärungspotenzial des Einzeleffekts von GeR-Gesamtniveau auf die Korrektheit im Präpositionsgebrauch annehmen.

Die eingangs formulierten Hypothesen (1./2.) zur ersten Leitfrage können den Daten nach somit nicht vollständig bestätigt werden: Anders als angenommen, zeigt sich eine ähnliche Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch in beiden L1-Subkorpora, und zwar sowohl auf dem Niveau der elementaren Sprachverwendung A2 als auch auf dem fortgeschrittenen GeR-Gesamtniveau B1+. Ein deutlicher Unterschied in der Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch lässt sich lediglich auf dem GeR-Gesamtniveau B2 beobachten, und zwar dergestalt, dass Lernende mit L1TV eine niedrigere Wahrscheinlichkeit für Korrektheit aufweisen als Lernende mit L1TA. Insgesamt nehmen die Wahrscheinlichkeiten für Korrektheit im Präpositionsgebrauch jedoch mit steigendem GeR-Gesamtniveau unabhängig von der L1 zu, wie eingangs angenommen wurde (Hypothese 3). Grundsätzlich führt eine vom Deutschen typologisch verschiedene Erstsprache (L1TV) den Daten nach nicht zu mehr Gebrauchsschwierigkeiten bei Präpositionen als eine zum Deutschen typologisch ähnliche Erstsprache (L1TA). Ein allgemeiner signifikanter Einfluss der Variable *Erstsprache der Lernenden* auf die Variable *Korrektheit* konnte nicht ermittelt werden. Der deutliche Unterschied auf dem GeR-Gesamtniveau B2 könnte jedoch als Hinweis auf ein unterschiedliches Erwerbtempo gesehen werden: Lernende mit L1TV brauchen im Vergleich zu Lernenden mit L1TA eine längere Zeit, um Präpositionen zielsprachlich angemessen zu realisieren. Dafür spricht auch die Beobachtung, dass für Lernende mit L1TV die Korrektheit im Präpositionsgebrauch mit steigendem GeR-Gesamtniveau zwar zunimmt, zwischen B1+ und B2 jedoch zu stagnieren bzw. leicht abzunehmen scheint während Lernende mit L1TA einen stetigen Anstieg in der Wahrscheinlichkeit für Korrektheit zeigen. Im Folgenden werden die Fehlertypen im Präpositionsgebrauch je L1-Subkorpus näher untersucht.

8.5.2.2 Auswertung Fehlertypen im Präpositionsgebrauch

Für die Auswertung der Fehlertypen im Präpositionsgebrauch sind ausschließlich die inkorrekten PP-Kontexte bei L1TV und L1TA relevant. Die statistischen Auswertungen der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse liefern in Bezug auf die Fehlertypen im Präpositionsgebrauch die folgenden zentralen Ergebnisse:

- 1) Beide Erstsprachengruppen L1TV und L1TA zeigen einen signifikanten Effekt des L1-Subkorpus auf den Fehlertyp im Präpositionsgebrauch: Lernende mit L1TA zeigen signifikant höhere Wahrscheinlichkeiten für realisierungsbedingte Abweichungen, d. h. sie lassen Präpositionen entweder aus oder fügen diese in Kontexten hinzu, in denen zielsprachlich keine Präposition gefordert ist. Lernende mit L1TA dahingegen weisen eine signifikant höhere Wahrscheinlichkeit für die inkorrekte Wahl von Präpositionen auf.
- 2) Mit steigendem GeR-Gesamtniveau nimmt die Wahrscheinlichkeit für realisierungsbedingte Fehler im Präpositionsgebrauch für L1TV und L1TA ab; die Wahrscheinlichkeit für wahlbedingte Fehler nimmt dahingegen zu.

Die Anzahl inkorrektter PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau habe ich bereits in Tabelle 36 präsentiert. In der folgenden Tabelle (Tab. 39) sind die inkorrekten PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau erneut wiedergegeben, nun mit der zusätzlichen Information zur Anzahl unterschiedlicher Lernertexte, d. h. Lernender, die im Datensatz der inkorrekten PP-Kontexte repräsentiert sind. Insgesamt sind im Datensatz inkorrektter PP-Kontexte 60 DaF-Lernende repräsentiert (s. Tab. 39). Hier wird nochmals deutlich: Die Datengrundlage ist sehr eingeschränkt. Die vorliegende Fallstudie kann als Pilotstudie angesehen werden, an der weiterführende Untersuchungen in Lernerkorpora anknüpfen können.

L1-Subkorpus	GeR-Gesamtniveau						AL Gesamt
	A2		B1+		B2		
	Inkorrekte PP-Kontexte	AL	Inkorrekte PP-Kontexte	AL	Inkorrekte PP-Kontexte	AL	
L1TV	24	14	18	7	43	17	38
L1TA	13	7	11	4	17	11	22

Tabelle 39: Inkorrekte PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau: Überblick. AL=Anzahl Lernertexte.

Zur Ermittlung und zum Vergleich einzelner Fehlertypen im Präpositionsgebrauch je L1-Subkorpus beziehe ich mich auf die in MERLIN vorliegenden Fehlerannotationen bzw. die dazugehörigen Fehlerannotationskategorien. Die entsprechenden Fehlerannotationen im Präpositionsgebrauch wurden im Rahmen der Datenerhebung aus dem Korpus extrahiert, manuell korrigiert (Näheres s. oben) und die einzelnen Fehlerannotationskategorien, die die einzelnen Fehlertypen im Präpositionsgebrauch repräsentieren, wurden ausgewertet. Die folgende Tabelle (Tab. 40) gibt zunächst einen zusammenfassenden Überblick über die absolute und relative Verteilung dieser Fehlertypen je L1-Subkorpus. In der Tabelle sind lediglich die GeR-Niveaus A2, B1+ und B2 repräsentiert, da lediglich diese GeR-Gesamtniveaus in der vorliegenden Fallstudie berücksichtigt werden (s. auch 8.5.2.1).

L1-Subkorpus	Fehlertyp im Präpositionsgebrauch				
	Wahl	Hinzufügung	Auslassung	Position	Summe
L1TV	32 (38%)	19 (22%)	34 (40%)	0 (0%)	85 (100%)
L1TA	28 (68%)	4 (10%)	9 (22%)	0 (0%)	41 (100%)

Tabelle 40: Fehlertypen im Präpositionsgebrauch je L1-Subkorpus (GeR-Gesamtniveaus A2, B1+ und B2): Überblick.

Tabelle 40 zeigt, dass die Anteile der Fehlertypen je L1-Subkorpus unterschiedlich hoch sind. Gemein ist den L1-Subkorpora, dass der Fehlertyp *Position* nicht vertreten ist. Dies stellt auf den ersten Blick ein auffälliges Ergebnis dar, da beide Sprachen des Subkorpus L1TV (Türkisch und Ungarisch) zwar keine Präpositionen aufweisen, jedoch über Postpositionen verfügen, mit denen u.a. räumliche Relationen realisiert werden können. Man könnte daher annehmen, dass Lernende mit L1TV die positionellen Eigenschaften dieser Einheiten auf die deutsche Sprache übertragen und deutsche Präpositionen in Poststellung realisieren. Dieses Phänomen

ist in der entsprechenden Forschungsliteratur bei erwachsenen Lernenden auch vereinzelt belegt (Meyer-Ingwersen 1975: 76, Keim 1984: 143, s. auch Kap. 4.1). In den hier vorliegenden Daten finden sich keine Hinweise auf eine derartige Interferenzerscheinung; zu ähnlichen Ergebnissen kommen z. B. auch Gutzmann/Turgay (2011) und Kuhberg (1990). Dennoch wäre ein Einfluss postpositionaler Einheiten aus der L1 für den Gebrauch von Präpositionen in der L2 Deutsch denkbar. Wie Bryant (2012) in ihrer Untersuchung zum kindlichen DaZ-Erwerb von Lokalisierungsausdrücken beobachten kann, verwenden Lernende mit L1 Türkisch in einem bestimmten Erwerbsstadium zwar Präpositionen, verknüpfen diese jedoch mit einem Präpositionaladverb, das als lokal spezifische Postposition dient (Beispiel: *Der Brief is bein Umschlag drin*) (Bryant 2012: 227). Inwiefern solch eine Strategie ebenfalls bei Lernenden mit L1TV im vorliegenden Datensatz aus MERLIN beobachtet werden kann, müsste in einer umfangreicheren Datenerhebung und -analyse unter Berücksichtigung von Präpositionaladverbien untersucht werden.

Im vorliegenden Datensatz sind damit drei verschiedene Fehlertypen im Präpositionsgebrauch repräsentiert. Die Fehlertypen *Auslassung* und *Hinzufügung* können zu einem übergeordneten Fehlertyp *Realisierung* zusammengefasst werden, da sowohl ausgelassene als auch inkorrekt hinzugefügte Präpositionen auf Schwierigkeiten hindeuten, die die Gebrauchskontexte von Präpositionen und PP betreffen. Der Fehlertyp Realisierung steht somit stellvertretend für die Unsicherheiten, die sich bei der Realisierung von Präpositionen zeigen. Während man beim Fehlertyp *Wahl*, d. h. bei der inkorrekten Wahl einer Präposition annehmen kann, dass Lernende zumindest über das sprachliche Wissen verfügen, dass eine Präposition in einem bestimmten syntaktischen Kontext zielsprachlich gefordert ist, lässt der Fehlertyp Realisierung eher vermuten, dass dieses Wissen (noch) nicht verfügbar ist. Die unteren Beispiele illustrieren den Fehlertyp Realisierung (1./2.) und den Fehlertyp Wahl (3./4.)²²⁹.

1. *Ich muss jeden Tag Deutsch lernen, weil ich 9. März eine Prüfung habe.* (L1TV, ID: 1061_0120346)
ZH1: Ich muss jeden Tag Deutsch lernen, weil ich am 9. März eine Prüfung habe.
2. *kann ich die Schlüssel zu dir geben* (ID: 1091_0000220)
ZH1: ich kann dir die Schlüssel geben
3. *Ich freue mich um Ihrer Angebot und die Möglichkeit für ein Vorstellungsgespräch.* (ID: 1023_0101841)
ZH1: Ich freue mich über Ihr Angebot und die Möglichkeit für ein Vorstellungsgespräch.
4. *Wir sollen um 8 Uhr zu Zentralplatz treffen.* (ID: 1071_0024817)
ZH1: Wir sollen uns um 8 Uhr am Zentralplatz treffen.

²²⁹ An dieser Stelle und bei weiteren Beispielen aus dem Korpus wird die Zielhypothese (ZH), so wie sie im Korpus vorliegt, zusätzlich angegeben, da die Fehlerannotationen sich auf die entsprechenden ZH beziehen.

Die untere Tabelle (Tab. 41) gibt die Verteilung der Fehlertypen *Realisierung* (R) und *Wahl* (W) je L1-Subkorpus und GeR-Gesamtniveau wieder. Insgesamt sind im Datensatz fehlerannotierter Präpositionen bzw. PP 60 Lernende repräsentiert (38 in L1TV und 22 in L1TA), die Verteilung der im Datensatz repräsentierten Lernertexte, d. h. Lernenden je GeR-Gesamtniveau kann Tabelle 39 (weiter oben) entnommen werden.

L1-Subkorpus	GeR-Gesamtniveau						Gesamt
	A2		B1+		B2		
	W	R	W	R	W	R	W/R
L1TV	6 (25%)	18 (75%)	6 (33%)	12 (67%)	20 (47%)	23 (53%)	85 (100%)
L1TA	6 (46%)	7 (54%)	10 (91%)	1 (9%)	12 (71%)	5 (19%)	41 (100%)

Tabelle 41: Fehlertypen Wahl (W) und Realisierung (R) je L1-Subkorpus und GeR-Gesamtniveau.

An dieser Stelle möchte ich darauf hinweisen, dass für die Auswertung zum Fehlertyp Realisierung lediglich die GeR-Gesamtniveaus A2 und B2 berücksichtigt werden, da im GeR-Gesamtniveau B1+ im L1-Subkorpus L1TA lediglich eine einzige Instanz dieses Fehlertyps vorliegt (s. Tab. 41). Die folgenden Auswertungen können als eine Pilotstudie betrachtet werden, an die anschließende Untersuchungen in deutschen Lernerkorpora anknüpfen können.

Die inkorrekt gebrauchten Präpositionen bzw. PP werden statistisch ausgewertet. Äquivalent zur Auswertung der Korrektheit im Präpositionsgebrauch nutze ich Regressionsanalysen und berechne je Fehlertyp gemischte Modelle (Näheres s. Kap. 7.2). Als Zielvariable (abhängige Variable) wähle ich jeweils den entsprechenden Fehlertyp (Wahl o. Realisierung). Die *Erstsprache der Lernenden* bzw. das L1-Subkorpus (L1.Zf) sowie das *GeR-Gesamtniveau* (Kompetenzniveau) werden als unabhängige Variablen ins jeweilige Modell integriert und stellen die festen Faktoren dar. Der/die einzelne Lernende wird als Zufallsfaktor (1|Lerner) festgelegt, um den individuellen Einfluss des/der Lernenden herauszurechnen. Mit den Regressionsmodellen wird überprüft, ob ein Effekt der L1 sowie des GeR-Gesamtniveaus auf die Fehlertypen im Präpositionsgebrauch, Realisierung und Wahl, ermittelt werden kann. In den Modellen werden ausschließlich die inkorrekten PP-Kontexte berücksichtigt (in den Regressionsmodellen sind diese jeweils mit *dat.subset* repräsentiert).

Zunächst möchte ich die Auswertungen zum Fehlertyp Wahl vorstellen. Das entsprechende erste Regressionsmodell (RM2) und sein Output sind unten angegeben. In diesem Regressionsmodell wird u.a. zunächst überprüft, ob eine Interaktion zwischen den festen Faktoren vorliegt.

RM2:

```
glmer(Wahl ~ L1.Zf*Kompetenzniveau + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.15415	0.55634	-0.277	0.7817
L1.ZfL1TV	-0.94446	0.72920	-1.295	0.1952
KompetenzniveauB1+	2.45674	1.18722	2.069	0.0385
KompetenzniveauB2	1.02962	0.76996	1.337	0.1811
L1.ZfL1TV:KompetenzniveauB1+	-2.05127	1.37175	-1.495	0.1348
L1.ZfL1TV:KompetenzniveauB2	-0.07077	0.95317	-0.074	0.9408

Die folgende Abbildung (Abb. 34) zeigt die Visualisierung des Modell-Outputs. Dabei sind die Wahrscheinlichkeiten für den Fehlertyp Wahl auf der Y-Achse abgebildet, die einzelnen GeR-Gesamtniveaus sind auf der X-Achse repräsentiert. Lernende mit L1TA sind als durchgezogene Linie repräsentiert, Lernende mit L1TV als gestrichelte Linie.

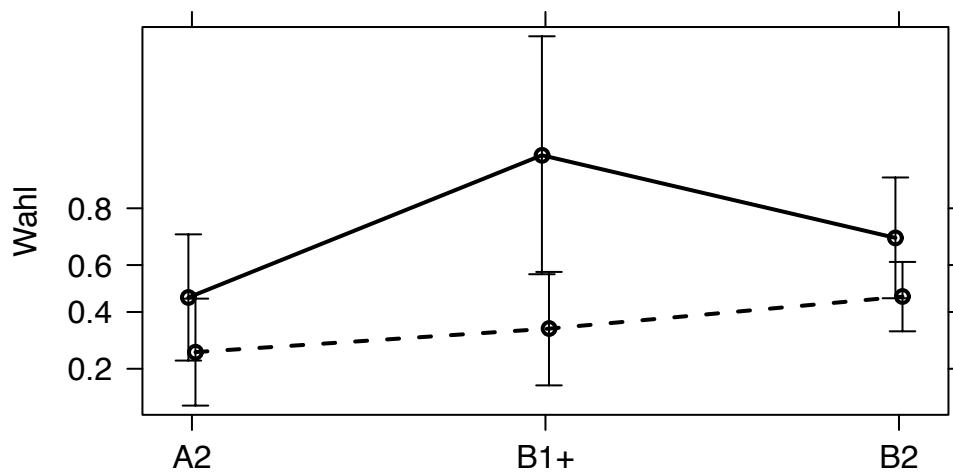


Abbildung 34: Fehlertyp Wahl im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM2. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.

Wie man Abbildung 34 entnehmen kann, zeigen Lernende mit L1TA in jedem GeR-Gesamtniveau eine höhere Wahrscheinlichkeit für wahlbedingte Fehler als Lernende mit L1TV. Der größte Unterschied zeigt sich dabei im GeR-Gesamtniveau B1+: Lernende mit L1TA zeigen hier eine deutlich höhere Wahrscheinlichkeit als Lernende mit L1TV: Die Standardabweichungen überlappen nicht die jeweiligen Mittelwerte, auch wenn hier bei Lernenden mit L1TA eine große Varianz vorliegt (s. die Standardabweichung). Auf dem GeR-Gesamtniveau A2 und B2 liegen die Wahrscheinlichkeiten für den Fehlertyp Wahl in beiden L1-Subkorpora dahingegen näher beieinander, hier gibt es jeweils entsprechende Überlappungen der Markierungen der Standardabweichungen. Für beide L1-Subkorpora lässt sich insgesamt ein Anstieg der Wahrscheinlichkeit wahlbedingter Fehler mit steigendem GeR-Gesamtniveau ausmachen: Bei Lernenden mit L1V zeigt sich ein stetiger Anstieg (von 20% über ca. 30% bis etwas über 40%), bei Lernenden mit L1TA dahingegen zunächst ein deutlicher Anstieg (von etwas mehr als 40% zu über 80%), und dann wieder eine leichte Abnahme (auf ca. 70%).

Das Ergebnis des Regressionsmodells RM2 zeigt keine Interaktion zwischen dem L1-Subkorpus und GeR-Gesamtniveau ($p = 0.275601$), d. h. die Variablen beeinflussen sich nicht

gegenseitig in ihren Effekten auf die Wahrscheinlichkeit des Fehlertyps Wahl. In einem modifizierten Modell (RM2a) werden daher lediglich die Einzeleffekte beider Variablen (L1-Subkorporus und GeR-Gesamtniveau) berücksichtigt. Das modifizierte Modell (RM2a) ist mit seinem Output im Folgenden angegeben:

RM2a:

```
glmer(Wahl ~ L1.Zf + Kompetenzniveau + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.09572	0.43823	0.218	0.82710
L1.ZfL1TV	-1.38358	0.42243	-3.275	0.00106
KompetenzniveauB1+	1.00656	0.54236	1.856	0.06347
KompetenzniveauB2	1.06070	0.46714	2.271	0.02317

Der Modell-Output zeigt einen signifikanten Effekt der L1 auf den Fehlertyp im Präpositionsgebrauch sowie einen signifikanten Effekt des GeR-Gesamtniveaus B2 im Vergleich zum Referenzniveau A2. Das Modell bzw. die Einzeleffekte sind in Abbildung 35 illustriert.

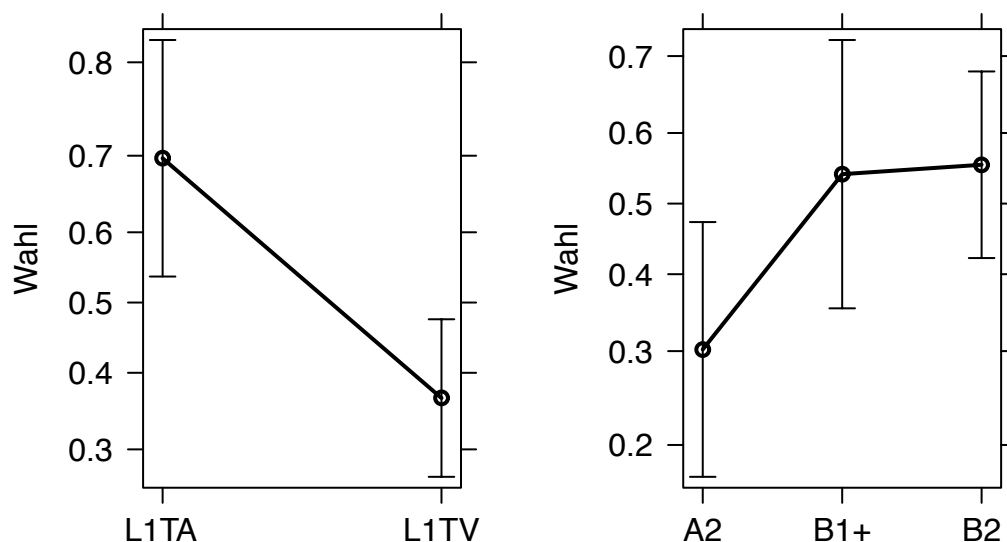


Abbildung 35: Fehlertyp Wahl im Präpositionsgebrauch je L1-Subkorporus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM2a.

Wie man Abbildung 35 entnehmen kann, zeigen Lernende mit L1TV eine signifikant niedrigere Wahrscheinlichkeit für wahlbedingte Präpositionsfehler als Lernende mit L1TA (etwas weniger als 40% vs. ca. 70%). Der Fehlertyp Wahl scheint daher charakteristisch für Lernende mit L1TA zu sein. Die typologische Nähe der L1 zur L2 Deutsch hat einen signifikanten Effekt auf die Wahrscheinlichkeit wahlbedingter Fehler im Präpositionsgebrauch. Dieses Ergebnis zeigt sich z. B. auch bei Pfaff (1984). Sie deutet wahlbedingte Fehler als Hinweis auf einen lexikalischen Transfer aus der L1 (Pfaff 1984: 296). Demnach lassen sich wahlbedingte Fehler im

Präpositionsgebrauch vorwiegend dann beobachten, wenn den Lernenden aus ihrer L1 Präpositionen und PP bekannt sind. Bedingt durch unterschiedliche semantische Reichweiten von Präpositionen in L1 und L2 kann dieses Wissen jedoch zu einem negativen Transfer, und damit zur inkorrekten Wahl von Präpositionen führen. Die typologische Nähe zwischen L1 (Englisch und Italienisch) und L2 Deutsch kann somit auch bestimmte Fehler im Sprachgebrauch begünstigen (vgl. Blumenthal-Dramé/Kortmann 2013: 292). Für Lernende mit L1TA zeigt sich die Unsicherheit im Gebrauch von Präpositionen und PP demnach hauptsächlich in der Wahl der korrekten Präposition. Beispiele für den wahlbedingte Fehler bei Lernenden mit L1TA sind unten angegeben, die Identifikationsnummer (ID) und die L1 sind jeweils in Klammern angegeben.

1. *Ich freue mich um Ihrer Angebot* (ID: 1023_0101841) (L1E)
ZH1: Ich freue mich über Ihr Angebot
2. *[...] ich bin nach meine " neue " Haus gefahren.* (ID: 1023_0108811) (L1I)
ZH1: [...] ich bin zu meinem " neuen " Haus gefahren.
3. *[...] ich bin für diese Stelle sehr interessiert.* (ID: 1023_0109717) (L1I)
ZH1: [...] ich bin an dieser Stelle sehr interessiert.
4. *Nach dem Abendessen sind wir in Zentralplatz [...] gegangen* (ID: 1061_0120882) (L1I)
ZH1: Nach dem Abendessen sind wir auf den Zentralplatz [...] gegangen
5. *Wenn mann auf China geht [...]* (ID: 1031_0003327) (L1I)
ZH1: Wenn man nach China geht [...]
6. *Wir sollen um 8 Uhr zu Zentralplatz treffen.* (ID: 1071_0024817) (L1E)
ZH1: Wir sollen uns um 8 Uhr am Zentralplatz treffen.
7. *Wenn ich in Deutschland gekommen war, [...]* (ID: 1023_0108814) (L1I)
ZH1: Als ich nach Deutschland gekommen war, [...]

Neben der allgemeinen hohen Wahrscheinlichkeit des Fehlertyps Wahl bei Lernenden mit L1TA zeigt Abbildung 35 ebenfalls, dass der Anteil dieses Fehlertyps mit Anstieg des GeR-Gesamtniveaus für beide L1-Subkorpora zunimmt. Ein deutlicher Anstieg zeigt sich zwischen dem GeR-Niveau der elementaren Sprachverwendung A2 und dem unteren Niveau der selbstständigen Sprachverwendung B1+ (von 30% zu 50%). Zwischen den GeR-Gesamtniveaus B1+ und B2 dahingegen gibt fast keinen Anstieg. Der Anstieg der Wahrscheinlichkeit wahlbedingter Fehler mit steigendem GeR-Gesamtniveau zeigt sich unabhängig vom L1-Subkorporus. Im Falle von Lernenden mit L1TV zeigt die steigende Wahrscheinlichkeit dieses Fehlertyps einen Lernfortschritt an: Wie man den Werten in Tabelle 41 entnehmen kann,

zeigen Lernende mit L1TV insgesamt einen größeren Anteil des Fehlertyps Realisierung (im Vergleich zu Lernenden mit L1TA). Der Anstieg des Fehlertyps Wahl bedeutet für diese Lernenden nun das Folgende: Präpositionen und PP werden nun in zielsprachlich angemessenen Kontexten realisiert, bei der Wahl der Präpositionen bestehen jedoch (noch) Unsicherheiten. Für Lernende mit L1TA zeigt der Anstieg der Wahrscheinlichkeit des Fehlertyps Wahl dahingegen keinen sichtbaren Lernfortschritt an: Wie man Abbildung 35 nämlich entnehmen kann, zeigen Lernende mit L1TA insgesamt ohnehin eine vergleichsweise hohe Wahrscheinlichkeit für diesen Fehlertyp. Der Anstieg des Fehlertyps Wahl mit zunehmendem GeR-Gesamtniveau illustriert folglich für Lernende mit L1TA, dass bei diesen Lernenden Unsicherheiten in der Wahl von Präpositionen bestehen bleiben.

Die Auswertung zum allgemeinen Einfluss des GeR-Gesamtniveaus auf den Fehlertyp Wahl zeigt jedoch (knapp) keinen signifikanten allgemeinen Effekt des GeR-Gesamtniveaus ($p=0.059143$). Das bedeutet, dass hier vor allem für die L1 ein Effekt auf den Fehlertyp Wahl im Präpositionsgebrauch ermittelt werden kann. Die Evaluation des Modells RM2a zeigt, dass das Modell die Varianz im vorliegenden Datensatz zu ca. 13,5% ($R^2=0.135238$) erklären kann. Das bedeutet, dass die Varianz in der Datenverteilung zum Fehlertyp Wahl zu 13,5% anhand des Einzeleffekts der L1 erklärt werden kann.

Im Folgenden präsentiere ich die Auswertung zum Fehlertyp *Realisierung*. Wie oben angemerkt, werden für die statistische Auswertung dieses Fehlertyps lediglich die GeR-Gesamtniveaus A2 und B2 ausgewertet, da im vorliegenden Datensatz im GeR-Gesamtniveau B1+ lediglich eine einzige Instanz des Fehlertyps Realisierung vorliegt (s. Tab. 41). Das Regressionsmodell (RM3²³⁰) für den Fehlertyp Realisierung entspricht dem Muster des Regressionsmodells zum Fehlertyp Wahl (s. oben) und ist im Folgenden mit seinem Output angegeben. Im diesem ersten Regressionsmodell zum Fehlertyp (RM3) wird ebenfalls eine Interaktion zwischen den festen Faktoren ins Modell integriert.

RM3:

```
glmer(Realisierung ~ Kompetenzniveau*L1.Zf + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.15415	0.55635	0.277	0.782
KompetenzniveauB2	-1.02962	0.76997	-1.337	0.181
L1.ZfL1TV	0.94446	0.72921	1.295	0.195
KompetenzniveauB2:L1.ZfL1TV	0.07077	0.95318	0.074	0.941

Das Ergebnis des Regressionsmodells ist in der folgenden Abbildung (Abb. 36) visualisiert: Lernende mit L1TV sind mit einer gestrichelten Linie repräsentiert, Lernende mit L1TA mit einer durchgezogenen Linie. Die Y-Achse illustriert die Wahrscheinlichkeit für den Fehlertyp

²³⁰ Im Regressionsmodell RM3 wurde die Reihenfolge der unabhängigen Variablen *L1.Zf* und *Kompetenzniveau* im Vergleich zu den anderen hier berichteten Modellen umgekehrt. Dies hat lediglich Auswirkungen auf die Visualisierung des Modell-Outputs und hat – wie ich überprüft habe – keine Auswirkungen auf das Ergebnis des Modell-Outputs.

Realisierung, die beiden GeR-Gesamtniveaus sind auf der X-Achse angegeben. Das Ergebnis des Regressionsmodells RM3 stellt ein Spiegelbild der entsprechenden Werte zum GeR-Gesamtniveau A2 und B2 des Modells zum Fehlertyp Wahl (RM2, s.o.) dar. Dies liegt daran, dass im ausgewerteten Datensatz ausschließlich die Fehlertypen Realisierung und Wahl enthalten sind (vgl. Tab. 41). Ergebnisse für den Fehlertyp Realisierung zeigen somit ‚die andere Seite‘ der Fehlertypverteilung in den GeR-Gesamtniveaus A2 und B2.

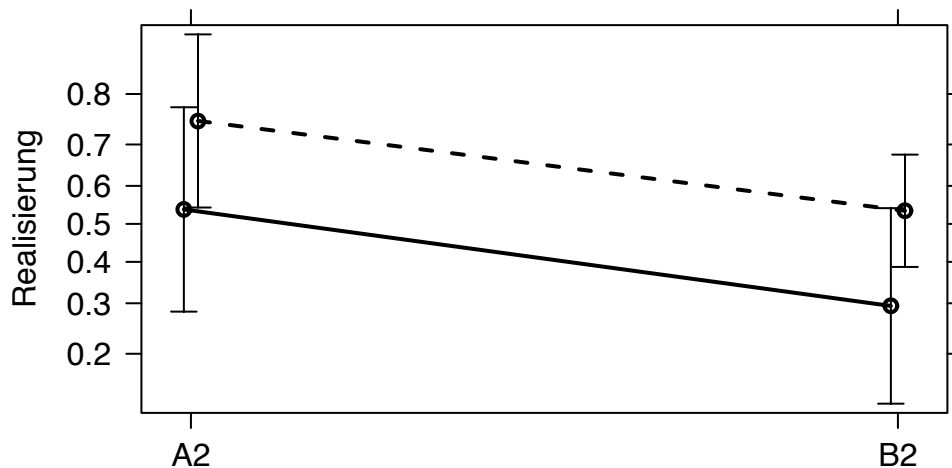


Abbildung 36: Fehlertyp Realisierung im Präpositionsgebrauch je L1-Subkorporum und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM3. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.

Abbildung 36 zeigt, dass Lernende mit L1TV eine konstant höhere Wahrscheinlichkeit für den Fehlertyp Realisierung aufweisen als Lernende mit L1TA. Die entsprechenden Wahrscheinlichkeiten unterscheiden sich für beide L1-Gruppen jedoch nicht stark voneinander. Weiterhin kann man Abbildung 36 entnehmen, dass bei Lernenden mit L1TV sowie Lernenden mit L1TA die Wahrscheinlichkeiten für den Fehlertyp Realisierung mit zunehmendem GeR-Gesamtniveau leicht sinken (auch wenn dies nicht abschließend klar ist, da entsprechende Werte für die dazwischen liegenden GeR-Gesamtniveaus fehlen). Das Ergebnis des Regressionsmodells zeigt keine Interaktion zwischen dem L1-Subkorporum und dem GeR-Gesamtniveau ($p=0.941$, s.o.). Das bedeutet, dass sich diese Variablen im Hinblick auf die Wahrscheinlichkeit für den Fehlertyp Realisierung nicht gegenseitig beeinflussen.

Ein entsprechend modifiziertes Modell mit ausschließlicher Berücksichtigung der Einzeleffekte beider Variablen (RM3a, s. unten) ist mit seinem Output im Folgenden angegeben. Die Visualisierung des Modell-Outputs zu RM3a findet sich in Abbildung 37.

RM3a:

```
glmer(Realisierung ~ L1.Zf + Kompetenzniveau + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.1300	0.4514	0.288	0.7733
L1.ZfL1TV	0.9860	0.4689	2.103	0.0355
KompetenzniveauB2	-0.9835	0.4544	-2.165	0.0304

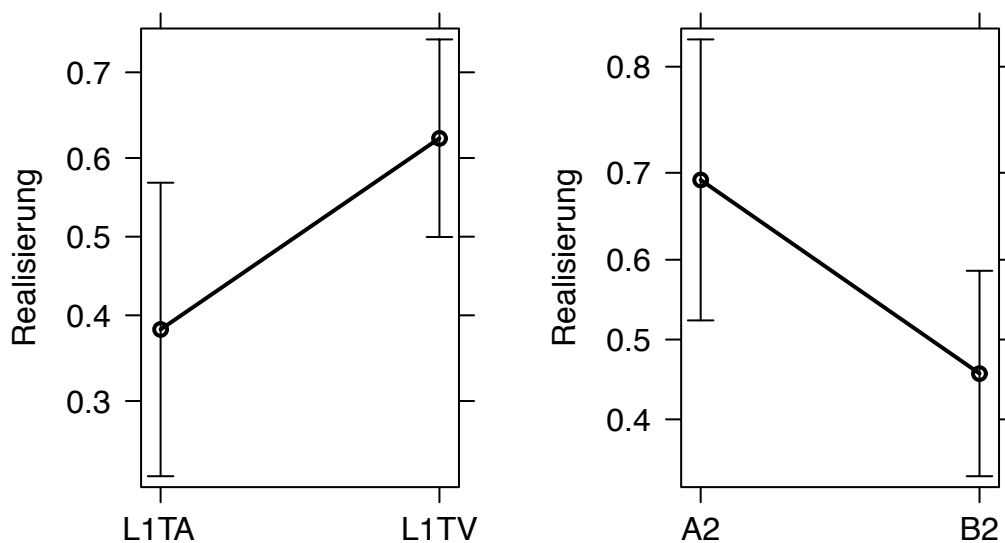


Abbildung 37: Fehlertyp Realisierung im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM3a.

Abbildung 37 zeigt deutlich, dass Lernende mit L1TV insgesamt eine deutlich höhere Wahrscheinlichkeit für Unsicherheiten in Bezug auf die Realisierung von Präpositionen zeigen (etwas mehr als 60%) als Lernende mit L1TA (knapp 40%). Zusätzlich lässt sich feststellen, dass im vorliegenden Datensatz eine größere Varianz bei Lernenden mit L1TA vorliegt (s. die Standardabweichungen in Abb. 37), für Lernende mit L1TV lässt sich eine geringere Varianz in den Wahrscheinlichkeiten des Fehlertyps Realisierung ermitteln. Das Ergebnis des Regressionsmodells (RM3a) zeigt einen signifikanten Effekt des L1-Subkorpus auf den Fehlertyp Realisierung ($p=0.0355$, s.o.) sowie einen signifikanten Effekt des GeR-Gesamtniveaus ($p=0.0304$, s.o.). Das bedeutet, dass für beide Variablen ein Effekt auf die Wahrscheinlichkeit des Fehlertyps Realisierung feststellbar ist.

Die unteren Beispiele illustrieren den Fehlertyp Realisierung bei Lernenden mit L1TV; die jeweilige Identifikationsnummer (ID) und die L1 der Lernenden ist jeweils in Klammern angegeben.

1. *Suche ich gerne auf eine neue Wohnung* (ID: 1091_0000192) (L1T)
ZH1: ich suche gerne eine neue Wohnung
2. *Ich mag auch natürlich mit anderen Gebräuchen kennen lernen, [...]* (ID: 1031_0003211) (L1U)
ZH1: Ich mag auch natürlich andere Gebräuche kennenlernen, [...]
3. *ich empfehle für Alle [...]* (ID: 1031_0003231) (L1U)
ZH1: ich empfehle allen [...]
4. *Ich warte Dich.* (ID: 1071_0243591) (L1T)
ZH1: Ich warte auf Dich.

5. *kannst du bitte meine katzen aufpassen.* (ID: 1091_0000140) (L1T)
ZH1: kannst du bitte auf meine Katzen aufpassen.

6. *Ich muss jeden Tag Deutsch lernen, weil ich 9. März eine Prüfung habe* (ID: 1061_0120346) (L1T)
ZH1: Ich muss jeden Tag Deutsch lernen, weil ich am 9. März eine Prüfung habe.

7. *[...] mich die Gewohnheiten zu gewöhnen.* (1031_0003225) (L1U)
ZH1: [...] mich an die Gewohnheiten zu gewöhnen.

Wie oben erwähnt, werden neben Auslassungen auch inkorrekte Hinzufügungen von Präpositionen zum Fehlertyp Realisierung gezählt. Hinzufügungen von Präpositionen durch Lernende mit L1TV können auf eine Kompensationsstrategie seitens der Lernenden hindeuten: Die fehlende Kenntnis des Konzepts *Präpositionen* muss bei Lernenden mit L1TV im Prozess des Erwerb der L2 Deutsch a) grundlegend erworben werden und b) müssen die Lernenden die syntaktischen Verwendungskontexte von Präpositionen und PP beherrschen, um Präpositionen und PP zielsprachlich gebrauchen zu können. Möglicherweise greift bei Lernenden mit L1TV an dieser Stelle eine Art ‚Hyperkorrektur‘ im Bereich Präpositionsgebrauch, mit der Folge, dass Präpositionen auch dort verwendet werden, wo sie zielsprachlich nicht gefordert sind. Auf eine ähnliche Strategie verweist Hoffmann (2016: 612) im Gebrauch definiter Artikel durch türkischsprachige Deutschlernende. Da im Türkischen ein definitiver Artikel fehlt, zeigen Lernende oft Schwierigkeiten im Gebrauch dieser Artikel im Deutschen und greifen auf die Faustregel zurück, „im Zweifel eher einen [Artikel] zu setzen“. Eine ähnliche Strategie könnte auch beim Gebrauch deutscher Präpositionen greifen. Im Gegensatz dazu ist Lernenden mit L1TA das Konzept der Präpositionen sowie mögliche syntaktische Verwendungskontexte aus ihrer L1 bereits bekannt, sodass bei Ihnen insgesamt weniger realisierungsbedingte Abweichungen prognostiziert werden können (s. Abb. 37).

Weiterhin macht Abbildung 37 sowie die entsprechenden Werte im Modell-Output (s.o.) deutlich, dass mit Zunahme der Sprachkompetenz die realisierungsbedingten Abweichungen insgesamt signifikant abnehmen, obwohl je GeR-Gesamtniveau deutliche Varianzen feststellbar sind (siehe die jeweiligen Standardabweichungen). Dieses Ergebnis zeigt sich unabhängig von der typologischen Nähe der L1 zur L2. Auf dem Niveau der elementaren Sprachverwendung A2 zeigt sich eine hohe Wahrscheinlichkeit für realisierungsbedingte Fehler (70%) – hier lassen sich somit realisierungsbedingte Fehler im Präpositionsgebrauch erwarten, und das unabhängig von der L1 der Lernenden. Dahingegen liegt die Wahrscheinlichkeit für realisierungsbedingte Fehler im Präpositionsgebrauch auf dem GeR-Gesamtniveau B2 nur noch bei ca. 45%. Es zeigt sich somit eine Abnahme der Wahrscheinlichkeit für diesen Fehlertyp mit steigendem GeR-Gesamtniveau. Die Evaluation des Modells RM3a zeigt, dass das Modell die Varianz im vorliegenden Datensatz zumindest zu 9,2% erklären kann ($R^2 = 0.09205347$).

Zusammenfassung

Den Ergebnissen der Auswertungen nach kann die zweite Leitfrage wie folgt beantwortet werden:

2. Welche Fehlertypen im Bereich Präpositionsgebrauch zeigen sich bei Deutschlernenden mit typologisch ähnlicher L1 (L1TA) bzw. typologisch verschiedener L1(L1TV)? (Wie) unterscheiden sich diese?

Die Auswertungen der computerunterstützten Fehleranalyse zeigen zwei im vorliegenden Datensatz vertretene Fehlertypen im Präpositionsgebrauch: Realisierung und Wahl, wobei zum Fehlertyp Realisierung sowohl Auslassungen von Präpositionen als auch inkorrekte Hinzufügungen von Präpositionen gezählt werden. Insgesamt zeigen die statistischen Auswertungen, dass für beide L1-Subkorpora jeweils spezifische Fehlertypen im Präpositionsgebrauch ermittelt werden können: Lernende mit einer vom Deutschen typologisch verschiedenen Erstsprache (L1TV) zeigen eine hohe Wahrscheinlichkeit für Unsicherheiten bei der Realisierung von Präpositionen, während Lernende mit einer zum Deutschen ähnlichen Erstsprache (L1TA) eine hohe Wahrscheinlichkeit für Probleme bei der Wahl der zielsprachlichen Präposition aufweisen. Diese Effekte sind unabhängig vom GeR-Gesamtniveau der Lernenden feststellbar. Weiterhin lässt sich unabhängig vom L1-Subkorpus feststellen, dass die Wahrscheinlichkeiten für Fehler in der Realisierung von Präpositionen mit zunehmendem GeR-Gesamtniveau signifikant sinken. Fehler in der Präpositionswahl dahingegen zeigen mit zunehmendem GeR-Gesamtniveau steigende Wahrscheinlichkeiten.

Die Hypothesen (4. 5. u. 6.) zur zweiten Leitfrage können den Daten nach folglich bestätigt werden: Lernende mit L1TV zeigen insgesamt eine höhere Wahrscheinlichkeit für Schwierigkeiten in der Realisierung von Präpositionen als Lernende mit L1TA, wohingegen diese wiederum eine höhere Wahrscheinlichkeit für wahlbedingte Fehler im Präpositionsgebrauch aufweisen. Für Lernende mit L1TV – für die insgesamt der dominante Fehlertyp Realisierung ermittelt werden konnte – lässt sich zudem ein Lernfortschritt ermitteln, der sich in der allgemeinen zunehmenden Wahrscheinlichkeit des Fehlertyps Wahl manifestiert. Dahingegen lässt sich ein derartiger Lernfortschritt für Lernende mit L1TA nicht ableiten – hier nimmt die Wahrscheinlichkeit für den ohnehin für Lernende mit L1TA dominanten Fehlertyp Wahl weiter zu. Insgesamt zeigen die Auswertungen, dass die typologische Distanz zwischen der L1 der Lernenden und der Zielsprache Deutsch sowie z.T. das GeR-Gesamtniveau einen Einfluss auf den Fehlertyp im Präpositionsgebrauch (Realisierung und Wahl) ausüben.

8.6 Zusammenfassung und Diskussion der Gesamtergebnisse

Die vorliegende Fallstudie widmete sich der Variable *Erstsprache (L1) der Lernenden* und ihrem Einfluss auf den Gebrauch von Präpositionen und PP im Deutschen. Ausgehend von der Annahme, dass die typologische Nähe der L1 der Lernenden zur Zielsprache Deutsch den Gebrauch von Präpositionen beeinflusst, wurden Lernende mit typologisch unterschiedlichen L1 betrachtet. Hierfür wurden entsprechende Subkorpora im Lernerkorpus MERLIN ausgewertet. Der Gebrauch von Präpositionen durch Lernende mit zum Deutschen typologisch verschiedenen L1 Türkisch (L1T) und Ungarisch (L1U) (Subkorpus L1TV) wurde

dem Gebrauch von Präpositionen durch Lernende mit dem Deutschen typologisch ähnlichen L1 Englisch (L1E) und Italienisch (L1I) (Subkorpus L1TA) gegenübergestellt. Die Untersuchung widmete sich ausschließlich der Korrektheit und den Fehlertypen im Präpositionsgebrauch. Hierfür wurden alle realisierten Präpositionen bzw. PP ausgewertet. Die im Korpus vorliegenden Fehlerannotationen im Bereich Präpositionsgebrauch wurden erhoben und manuell bereinigt. Im Rahmen einer computerunterstützten Fehleranalyse und einer kontrastiven Interlanguage-Analyse wurden die Korrektheit im Präpositionsgebrauch und die Fehlertypen im Präpositionsgebrauch ermittelt und die beiden L1-Subkorpora kontrastiert. Die Daten wurden mittels Regressionsanalysen (gemischte Modelle) ausgewertet. Es wurde überprüft, ob bzw. inwiefern die typologische Nähe der L1 zur Zielsprache Deutsch einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch hat. Für die Auswertungen wurde das GeR-Gesamtniveau der Lernenden berücksichtigt. Die Auswertungen liefern wertvolle Erkenntnisse in Bezug auf die Unsicherheiten im Gebrauch von Präpositionen und PP im Verlauf zunehmender Sprachkompetenz und zeigen z.T. spezifische Tendenzen bei Lernenden unterschiedlicher L1. In der vorliegenden Fallstudie wurden 644 PP-Kontexte (L1TV) bzw. 433 PP-Kontexte (L1TA) von insgesamt 88 DaF-Lernenden ausgewertet.

Die statistischen Auswertungen der Korrektheit im Präpositionsgebrauch zeigen, dass Lernende mit L1TV sowie Lernende mit L1TA ähnlich viele bzw. wenige Unsicherheiten im Hinblick auf den Gebrauch von Präpositionen und PP im Deutschen aufweisen: Hier kann insgesamt kein Effekt des L1-Subkorpus und damit kein Effekt der typologischen Nähe der L1 zur Zielsprache Deutsch ermittelt werden. Im Vergleich zu einer typologischen Distanz der L1 führt eine typologische Nähe der L1 zum Deutschen den Ergebnissen nach folglich nicht allgemein zu einer signifikant höheren Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch, anders als eingangs angenommen wurde (1./2. Hypothese). Entsprechende Tendenzen zur Unabhängigkeit der Schwierigkeiten im Präpositionsgebrauch von der L1 der Lernenden finden sich auch in anderen Untersuchungen im Kontext der Z/FSE-Forschung, z. B. bei Fries (1988), Gibson/Hufeisen/Libben (2001), Hufeisen/Gibson (2002) und Balci (2006) (vgl. Kap. 4.1).

Lediglich auf dem GeR-Gesamtniveau B2 lässt sich in den analysierten Daten ein (leicht) signifikanter Unterschied zwischen den L1-Subkorpora ermitteln: Hier zeigen Lernende mit L1TA eine deutlich höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch als Lernende mit L1TV. Dies kann eventuell als Indiz für einen Einfluss der typologischen Distanz auf das Erwerbtempo von Präpositionen und PP gesehen werden: Lernende mit L1TV benötigen im Vergleich zu Lernenden mit L1TA eine längere Zeit, um Präpositionen zielsprachlich angemessen zu realisieren. Damit zeigen die Ergebnisse der vorliegenden Fallstudie Ähnlichkeiten zu den Ergebnissen von Turgay (2010a) und Bryant (2012) (s. Kap. 4.1). Die Auswertungen bestätigen jedoch die 3. Hypothese: Unabhängig von der L1 der Lernenden nimmt die Korrektheit mit steigender Sprachkompetenz zu.

Lernende mit zum Deutschen typologisch ähnlichen bzw. typologisch verschiedenen L1 zeigen den Daten nach insgesamt zwar eine ähnliche Wahrscheinlichkeit für die Korrektheit im Gebrauch deutscher Präpositionen und PP, jedoch lassen sich charakteristische

Fehlertypen je L1-Subkorpus ermitteln. Für die Auswertung der Fehlertypen im Präpositionsgebrauch wurden ausschließlich die inkorrekten PP-Kontexte in den Subkorpora L1TV und L1TA berücksichtigt. Basierend auf der Auswertung der Fehlerannotationen lässt sich zunächst festhalten, dass Lernende mit L1TV sowie Lernende mit L1TA ähnliche Fehlertypen im Präpositionsgebrauch aufweisen, die Wahrscheinlichkeiten für die einzelnen Fehlertypen sind jedoch von der typologischen Nähe der L1 zur Zielsprache Deutsch beeinflusst: Lernende mit L1TV zeigen insgesamt eine signifikant höhere Wahrscheinlichkeit für realisierungsbedingte Fehler im Präpositionsgebrauch, d. h. sie lassen Präpositionen aus oder fügen Präpositionen in syntaktischen Kontexten hinzu, in denen zielsprachlich keine Präposition bzw. PP gefordert ist. Damit bestätigen die Ergebnisse die 4. Hypothese. Hinzufügungen von Präpositionen können als Hinweis auf eine Art Kompensationsstrategie gesehen werden: Auf Basis fehlenden (erst)sprachlichen Wissens um das Konzept der Präposition sowie über mögliche Gebrauchskontexte von Präpositionen und PP im Deutschen neigen Lernende mit L1TV vielerorts dazu, im Zweifelsfall eine Präposition zu realisieren (vgl. Hoffmann 2016: 612). Auslassungen von Präpositionen deuten ebenfalls auf Unsicherheiten in Bezug auf die Realisierungskontexte von Präpositionen hin und könnten auf einen Einfluss des erstsprachlichen Wissens hindeuten. Ähnliche Ergebnisse finden sich z.T. auch bei Viorel (1973), Bouillon (1977) Keim (1984) und Grießhaber (2007b) – auch wenn letzterer vor allem von wahlbedingten Fehlern bei Lernenden mit L1 Türkisch berichtet (s. Kap. 4.1). Ein interessanter Befund ist, dass im vorliegenden Datensatz keine positionsbedingten Präpositionsfehler zu beobachten sind: Lernende mit L1TV zeigen demnach keinen negativen Transfer PP-äquivalenter postpositionaler Elemente aus ihrer L1. Dieses Ergebnis entspricht auch den Beobachtungen bisheriger Untersuchungen (vgl. Gutzmann/Turgay 2011, Turgay 2010a, Kuhberg 1990, s. Kap. 4.1).

Für Lernende mit L1TA zeigt sich eine signifikant niedrige Wahrscheinlichkeit für realisierungsbedingte Fehler im Präpositionsgebrauch, jedoch kann für diese Gruppe insgesamt eine signifikant höhere Wahrscheinlichkeit für die inkorrekte Wahl von Präpositionen ermittelt werden, was somit die 5. Hypothese bestätigt. Dieser Befund hängt vermutlich damit zusammen, dass Lernenden mit L1TA das Konzept der Präposition sowie mögliche Gebrauchskontexte von Präpositionen und PP aus ihrer L1 bekannt sind. Lernende mit L1TA verfügen somit über das entsprechende sprachliche Wissen, wodurch es vergleichsweise seltener zu Fehlern in der Realisierung von zielsprachlich geforderten Präpositionen kommt, sondern eher z. B. zu Übergeneralisierungen im Gebrauch einzelner Präpositionen und damit zur inkorrekten Präpositionswahl. Ähnliche Befunde präsentiert auch Pfaff (1984). Die wahlbedingten Fehler bei Lernenden mit L1TA könnten ebenfalls auf Strategien von Hyperkorrekturen hindeuten, so wie Gibson/Hufeisen/Libben (2001) es in ihrer Untersuchung annehmen (vgl. Kap. 4.1). Um dies jedoch abschließend beurteilen zu können, wären weitere Analysen des Präpositionsgebrauchs bei Lernenden mit L1TA notwendig.

Neben dem Effekt der L1 auf den Fehlertyp im Präpositionsgebrauch kann zusätzlich ein Effekt des GeR-Gesamtniveaus ermittelt werden: Die Ergebnisse der statistischen Auswertung auf den GeR-Gesamtniveaus A2, B1+ und B2 bestätigen die 6. Hypothese, in dem sie zeigen, dass unabhängig von der typologischen Nähe der L1 zur Zielsprache Deutsch die

Wahrscheinlichkeit für realisierungsbedingte Fehler im Präpositionsgebrauch mit steigender Sprachkompetenz insgesamt abnimmt, während diese für wahlbedingte Fehler im Präpositionsgebrauch tendenziell insgesamt zunimmt. Von ähnlichen Ergebnissen berichten u.a. auch Meyer-Ingwersen (1975), Keim (1984), Apeltauer (1987), Lütke (2010) sowie Gutzmann/Turgay (2011) (s. Kap. 4.1). Diese Ergebnisse zur Entwicklung der Fehlertypen je GeR-Gesamtniveau zeigen: Der Zuwachs an sprachlichem Wissen, der sich im Anstieg der Sprachkompetenz manifestiert, geht einher mit einer Änderung des wahrscheinlichsten Fehlertyps: Vom Fehlertyp Realisierung (Auslassungen und Hinzufügungen von Präpositionen) zum Fehlertyp Wahl. Für Lernende mit L1TV zeigt diese Änderung einen deutlichen Lernfortschritt an. Der Fehlertyp Wahl lässt zumindest vermuten, dass das sprachliche Wissen um die Realisierungskontexte von Präpositionen und PP dahingehend ausgebaut wird, dass lediglich die Wahl der zielsprachlich korrekten Präposition eine Herausforderung darstellt, jedoch nicht mehr die Frage, ob die Realisierung einer Präposition überhaupt gefordert ist. Lernende mit L1TV, so könnte man sagen, ‚erarbeiten‘ sich somit mit steigendem Kompetenzniveau den Fehlertyp Wahl. Dies ist für Lernende mit L1TA anders: Diese zeigen insgesamt eine höhere Wahrscheinlichkeit für diesen Fehlertyp und bauen diese mit steigendem GeR-Gesamtniveau weiter aus.

Zusammenfassend lässt sich festhalten, dass die vorliegende Fallstudie wertvolle Erkenntnisse zum Einfluss der typologischen Nähe der L1 zur Zielsprache Deutsch auf die spezifischen Schwierigkeiten im Gebrauch von Präpositionen und PP liefert. Die Auswertung des vorliegenden (jedoch z. T. eingeschränkten) Datensatzes aus dem Lernerkorpus MERLIN zeigt, dass die Korrektheit im Präpositionsgebrauch insgesamt nicht anhand der typologischen Distanz der L1 zur Zielsprache Deutsch prognostiziert werden kann. Dahingegen können die Fehlertypen im Präpositionsgebrauch (Realisierung und Wahl) zu einem gewissen Grad durch die typologische Distanz der L1 vorhergesagt werden. Weitere Untersuchungen auf Basis einer umfangreicheren Datengrundlage wären notwendig, um diese Ergebnisse zu validieren.

9. Fallstudie 2: Syntaktische Funktion der PP

9.1 Einleitung

In der zweiten Fallstudie²³¹ im Lernerkorpus MERLIN wird die Variable *Syntaktische Funktion der PP* und ihr Einfluss auf den Gebrauch der Präposition innerhalb der PP untersucht. Im Vordergrund stehen adverbiale PP und Objekt-PP. In bisherigen Untersuchungen zum Erwerb und Gebrauch von Präpositionen durch DaF-Lernende fehlen systematische Analysen und Gegenüberstellungen des Gebrauchs von Präpositionen in unterschiedlichen syntaktischen Kontexten der entsprechenden PP. Die vorliegende Fallstudie soll einen Beitrag zur Schließung dieser Forschungslücke leisten. Adverbiale PP und Objekt-PP weisen unterschiedliche Form-

²³¹ In Weber (2018) wurden bereits vorläufige Ergebnisse dieser Fallstudie publiziert. Im Sammelband von Abel et al. (2019) werden die zentralen Ergebnisse dieser Fallstudie in abgewandelter Form publiziert (Weber 2019, i.Dr.) (Stand: Juni 2019).

Funktionsverbindungen auf, die einen Einfluss auf den Gebrauch dieser Strukturen und der darin enthaltenen Präpositionen durch DaF-Lernende erwarten lassen (s. Kap. 1.2).

9.2 Untersuchungsgegenstand und Leitfragen

Ziel der Fallstudie ist zu untersuchen, wie DaF-Lernende Präpositionen in PP unterschiedlicher syntaktischer Funktion gebrauchen und inwiefern die beobachtbaren Gebrauchsschwierigkeiten in Abhängigkeit zu der syntaktischen Funktion der PP stehen. Das sprachliche Kompetenzniveau der Lernenden wird dabei jeweils berücksichtigt. In Kapitel 1.2 wurde bereits dargelegt, dass PP unterschiedliche syntaktische Funktionen einnehmen können. In der vorliegenden Fallstudie stehen Objekt-PP und adverbiale PP im Vordergrund. Im Folgenden gehe ich kurz auf die Form-Funktionsverbindungen dieser beiden syntaktischen Funktionen von PP ein.

Objekt-PP und adverbiale PP unterscheiden sich in der Prototypikalität und Zuverlässigkeit ihrer Form-Funktionsverbindung. Zuerst kann festgestellt werden, dass im Deutschen PP keine prototypische Form von Verbobjekten repräsentieren. Prototypische Objekte werden im Deutschen i.d.R. durch reine Kasusobjekte in Form von Nominalphrasen (NP) realisiert (Duden 2016: 399, Eisenberg 2013: 64, Hundt 2001: 167, s. auch Kap. 1.2.1). Die Präferenz für NP für die Realisierung von Objekten wird auch in der Kasushierarchie (Eisenberg 2013: 63) sichtbar, in der Objekt-PP den untersten Rang einnehmen. Sie gelten damit als stark morphologisch markiert und im Vergleich zu anderen Kasusformen weniger frequent²³². Adverbiale (temporal, lokal, modal) werden dahingegen prototypisch in der Form von PP ausgedrückt und die adverbiale Funktion gilt auch als Hauptfunktion von PP (Heidolph/Flämig/Mosch 1981: 370). Dies bedeutet, dass PP prototypisch adverbiale PP darstellen und Adverbialien wiederum prototypisch in der Form von PP realisiert werden²³³.

Die Form-Funktionsverbindung *PP = Adverbiale* ist damit prototypischer und zuverlässiger als die Form-Funktionsverbindung *PP = Objekt*. Die folgende Tabelle (Tab. 42) fasst die Form-Funktionsverbindungen von Objekt-PP und adverbialer PP in einem Überblick zusammen.

²³² Hier sind vor allem die Kasus Nominativ, Akkusativ und Dativ gemeint. Eisenberg (2013: 63) platziert Objekt-PP und Genitiv-NP auf denselben Rang (s. Kap. 1.2.1).

²³³ Eine Abfrage in der Tübinger Baubank TüBa-D/Z (Version 10.1) zur Verteilung syntaktischer Funktionen von PP liefert hierfür weitere Evidenz: Adverbiale PP weisen eine höhere Frequenz auf als Objekt-PP. Eine Abfrage in der Baubank in Bezug auf die entsprechenden Kategorien (PP als *V-MOD* einerseits und PP als *OPP* bzw. *FOPP* andererseits, s. Telljohann et al. 2015: 142f.) lieferte folgendes Ergebnis: 54539 Treffer für adverbiale PP (Kategorie *V-MOD*) und 36828 Treffer für Objekt-PP (Kategorie *OPP* für obligatorische Präpositionalobjekte und Kategorie *FOPP* für fakultative Präpositionalobjekte). Das Verhältnis beider PP-Funktionen ist 1:1,5. Die Baubank TüBa-D/Z (Version 10.1) ist ein syntaktisch annotiertes Korpus, das auf Basis von Artikeln der Tageszeitung *die tageszeitung (taz)* erstellt wurde. Die Baubank umfasst insgesamt 3644 Zeitungsartikel und 95595 Sätze. Für weitere Informationen zur Baubank und den darin enthaltenen Annotationen s. Telljohann et al. (2015).

Form	Syntaktische Funktion	Form-Funktionsverbindung	Beispiele
PP	Adverbiale PP	prototypisch und zuverlässig	<i>Er öffnete die Flasche <u>auf der Bank</u>.</i> <i>Ich sitze <u>auf einer Bank</u>.</i>
	Objekt-PP	nicht prototypisch und weniger zuverlässig	<i>Sie wartete <u>auf ihren Vater</u>.</i> <i>Ich bin stolz <u>auf meine Tochter</u>.</i>

Tabelle 42: Objekt-PP und adverbiale PP: Form-Funktionsverbindungen.

Aktuelle Theorien zum L2-Erwerb gehen davon aus, dass die Prototypikalität und Zuverlässigkeit von Form-Funktionsverbindungen linguistischer Strukturen, auch „cue validity“ (Ellis 2003: 95) genannt, einen Einfluss auf den Erwerb dieser Strukturen ausübt (s. Kap. 2). Prototypische, d. h. zuverlässige Form-Funktionsverbindungen werden demnach schneller bzw. früher erworben als weniger prototypische bzw. weniger zuverlässige Form-Funktionsverbindungen (MacWhinney 2001: 70). Die zentralen Leitfragen der vorliegenden Fallstudie lauten:

1. Welche Unterschiede zeigen sich bei der Korrektheit im Präpositionsgebrauch innerhalb von PP unterschiedlicher syntaktischer Funktion (Objekt-PP/adverbiale PP)?
2. Welche Fehlertypen im Präpositionsgebrauch treten häufig auf und (wie) unterscheiden sich diese gemäß der syntaktischen Funktion der PP (Objekt-PP/adverbiale PP)?

Bei beiden Leitfragen wird jeweils zusätzlich das GeR-Gesamtniveau der Lernenden berücksichtigt.

Entsprechend den Annahmen aktueller, gebrauchsbasierter Ansätze zum L2-Erwerb (u.a. Ellis/Wulf 2015, Ellis 2003, MacWhinney 2001) (Kap. 2) gehe ich davon aus, dass Schwierigkeiten im Gebrauch von Präpositionen und PP durch die Prototypikalität und Zuverlässigkeit von Form-Funktionsverbindungen beeinflusst sind. Gebrauchsschwierigkeiten werden auch in dieser Fallstudie durch die Korrektheit und die Fehlertypen im Präpositionsgebrauch operationalisiert. In der vorliegenden Fallstudie gehe ich von folgenden Hypothesen aus, die ich im Rahmen der korpusgestützten Untersuchung statistisch überprüfe.

Hypothesen zu Leitfrage 1:

1. Adverbiale PP weisen insgesamt eine höhere Korrektheit im Präpositionsgebrauch auf als Objekt-PP.
2. Im Vergleich zu mittleren und höheren Niveaus der Sprachverwendung zeigen adverbiale PP auf dem Niveau der elementaren Sprachverwendung (A2 und A2+) eine deutlich höhere Korrektheit als Objekt-PP.
3. Unabhängig von der syntaktischen Funktion der PP nimmt die Korrektheit im Präpositionsgebrauch mit Anstieg des Kompetenzniveaus zu.

Hypothesen zu Leitfrage 2:

4. In Objekt-PP werden Präpositionen häufiger ausgelassen als in adverbialen PP.
5. In adverbialen PP werden Präpositionen häufiger inkorrekt gewählt als in Objekt-PP.
6. Unabhängig von der syntaktischen Funktion der PP nehmen Auslassungen von Präpositionen mit Anstieg des Kompetenzniveaus ab, während wahlbedingte Präpositionsfehler zunehmen.

9.3 Datenauswahl und Datenerhebung

Für die vorliegende Fallstudie in MERLIN sind vor allem die Metadaten der Lernertexte (Identifikationsnummer, GeR-Gesamtniveau), die Ebene der Zielhypothese (ZH1) sowie die Fehlerannotationskategorien im Bereich Präpositionsgebrauch von Relevanz.

Anders als in der ersten Fallstudie werden im Rahmen der vorliegenden Fallstudie nicht alle Präpositionen und PP im Lernerkorpus ausgewertet. Für diese Fallstudie werden zunächst primäre, lokale Präpositionen ausgewählt, da diese die prototypischen Vertreter der Wortartenklasse repräsentieren und in PP beider hier untersuchter syntaktischer Funktionen (Objekt-PP und adverbiale PP) auftreten können. Objekt-PP sind zudem auf diese diachron gesehen ältere Schicht von Präpositionen beschränkt²³⁴ (u. a. Duden 2016: 613f., Breindl 1989: 10, Breindl 2006: 938, GDS 1997: 1372, s. Kap. 1.1). Grundlage für die Datenauswahl stellt folgende Liste lokaler Präpositionen aus Duden (2016: 616) dar:

ab, an, auf, aus, außer, bei, bis, durch, gegen, hinter, in, nach, neben, um, unter, über, von, vor, zu, zwischen

Für diese Präpositionen werden in der lexikalischen Datenbank dlexDB (Heister et al. 2011)²³⁵ die Häufigkeiten ermittelt. Die Datenbasis von dlexDB bildet das Kernkorpus des Digitalen Wörterbuchs der deutschen Sprache (DWDS)²³⁶ – ein deutsches Referenzkorpus mit mehr als 100 Mio. Token, die zu ähnlichen Anteilen auf die Textsortenbereiche Belletristik, Zeitungsartikel, Wissenschaftstexte sowie Texte aus Gebrauchsliteratur verteilt und hinsichtlich der Dekaden des 20. Jahrhunderts ausgewogen sind (Geyken 2007). Die Ermittlung der Häufigkeiten obiger Präpositionen in dlexDB erfolgt anhand einer Listenabfrage der annotierten Types (vgl. Heister et al. 2011: 14), wobei die Groß- und Kleinschreibung gleichermaßen erfasst wird. Zusätzlich werden gemäß dem Stuttgart-Tübingen-Tagset (STTS) (Schiller et al. 1999) die Part-of-Speech-Kategorien APPR bzw. APPRART für die Abfrage spezifiziert. Die folgende Tabelle (Tab. 43) zeigt die sechs häufigsten Präpositionen in dlexDB²³⁷, Verschmelzungen²³⁸ sind dabei berücksichtigt. Die Groß- und Kleinschreibung ist in der Tabelle zusammengefasst.

²³⁴ Dies gilt bis auf die Präposition *ab*, die keine Objekt-PP bildet (Breindl 1989: 10).

²³⁵ dlexDB ist nach kostenloser Registrierung frei zugänglich unter <http://dlexdb.de/>.

²³⁶ Das Kernkorpus kann unter <https://www.dwds.de/> abgefragt werden.

²³⁷ Stand: Januar 2019.

²³⁸ Die entsprechenden Verschmelzungen lauten: *im, ins, zum, zur, aufs, am, beim*.

Rangfolge	Präposition (STTS-Kategorien APPR/APPRART)	Gesamthäufigkeit (absolut)
1	<i>in</i>	2 383678,00
2	<i>von</i>	1 145811,00
3	<i>zu</i>	775534,00
4	<i>auf</i>	672159,00
5	<i>an</i>	633637,00
6	<i>bei</i>	354238,00

Tabelle 43: Die sechs häufigsten Präpositionen (inkl. Verschmelzungen) in dlexDB.

Wie Tabelle 43 entnommen werden kann, weisen die abgefragten Präpositionen eine unterschiedliche Frequenz auf. Die Präposition *in* stellt die häufigste Präposition dar, gefolgt von den Präpositionen *von* und *zu*. Die Präpositionen *an* und *auf* weisen eine ähnliche Frequenz auf, unterscheiden sich aber deutlich von den Frequenzen ihrer direkten Nachbarn (*zu* und *bei*). Zwar kann auf Basis der ermittelten Werte in dlexDB nicht auf den sprachlichen Input der DaF-Lernenden, die in MERLIN repräsentiert sind, geschlossen werden. Da mir jedoch keine Informationen zum sprachlichen Input dieser Lernenden vorliegen, stellt die Abfrage in dlexDB einen empirischen Richtwert dar, an dem ich mich für die vorliegende Fallstudie orientiere.

Um angemessene Präpositionen für die hier vorliegende Fallstudie auszuwählen, waren weitere Überlegungen notwendig. Die Präposition *in* wird nur von wenigen Verben regiert, d.h. Objekt-PP mit der Präposition *in* sind (im Vergleich zu adverbialen PP) in der Zielsprache Deutsch eher selten anzutreffen²³⁹ und eignen sich daher nicht besonders gut für die vorliegende Untersuchung. Wie eine Abfrage und stichprobenartige Durchsicht im Lernerkorpus MERLIN aufzeigte, stellen ebenfalls die Präpositionen *von* und *zu* keine gut geeigneten Kandidaten für die vorliegende Fallstudie dar: Die Präposition *von* wird vorwiegend in attributiven Konstruktionen verwendet (anstelle einer Genitiv-NP), die Präposition *zu* taucht sehr oft als Flexionselement des zu-Infinitivs auf. Die Präpositionen *auf* und *an* dahingegen weisen einen vielfältigen Gebrauch auf; zudem werden sie von vielen deutschen Verben regiert (E-Valbu, PONS 2019). Dies – sowie die relativ hohe Frequenz dieser Präpositionen in der Zielsprache (vgl. Tab. 43) – macht die Präpositionen *auf* und *an* zu angemessenen Kandidaten für die vorliegende Fallstudie. Daher werden die Präpositionen *auf* und *an* ausgewählt. Die vorliegende Fallstudie beschränkt sich folglich auf den Einfluss der syntaktischen Funktion auf den Gebrauch von Objekt-PP und adverbialen PP mit den Präpositionen *an* und *auf*.

Die vorliegende Fallstudie hat zum Ziel, den Gebrauch dieser ausgewählten Präpositionen und PP durch DaF-Lernende zu untersuchen. Dabei bin ich vordergründig daran interessiert, wie Präpositionen und PP in der Zielsprache gebraucht werden. Somit stehen

²³⁹ Eine Durchsicht der Verben mit einer PP-Ergänzung in E-Valbu ergab, dass nur wenige Verben eine Präpositionsreaktion mit *in* aufweisen (E-Valbu). Ein weiterer Hinweis darauf, dass nur wenige deutsche Verben die Präposition *in* regieren, findet sich z. B. in DaF-Lehrmaterialien zum Thema Verben mit Präpositionen (PONS 2019).

zielsprachige Kontexte von PP im Fokus, die im Lernerkorpus auf Ebene der ZH1 repräsentiert sind. Durch eine Abfrage auf ZH1-Ebene erhalte ich sowohl PP-Kontexte, in denen die zielsprachliche Präposition bzw. PP korrekt realisiert wurde als auch solche PP-Kontexte, in denen zielsprachlich Korrekturen im Bereich des Präpositionsgebrauchs vorgenommen wurden. Eine Abfrage auf der Ebene der Wortartenannotationen liefert in MERLIN nur eine begrenzte Anzahl an PP-Kontexten, da diese Abfrage ausschließlich solche PP-Kontexte erfasst, in denen die Präposition als syntaktische Einheit von den Lernenden realisiert wurde und folglich als Präposition annotiert ist. PP-Kontexte, in denen der/die Lernende eine Präposition ausgelassen hat, können auf diese Weise jedoch nicht ermittelt werden. Eine Korpusabfrage auf der ZH1-Ebene ermöglicht es, sowohl die realisierten als auch die ausgelassenen PP-Kontexte im Korpus zu ermitteln. Aus diesem Grund erfolgt die Korpusabfrage auf der Ebene der ZH1, wobei die Groß- und Kleinschreibung gleichermaßen berücksichtigt werden. Da die ZH1-Ebene nicht mit Wortartenannotationen angereichert ist, muss die Korpusabfrage wortformbasiert erfolgen. Pseudotreffer wie beispielsweise abtrennbare Verbpartikeln (Beispiel: *Ruf mich bitte an.*), die mit dieser Art der Korpusabfrage nicht ausgeschlossen werden können, werden in der Phase der Datenaufbereitung manuell aussortiert.

Die Abfrage auf Ebene der ZH1 ermittelt PP-Kontexte, in denen die entsprechenden Präpositionen zielsprachlich syntaktisch gefordert sind. Das bedeutet, dass durch eine Abfrage auf Ebene der ZH1-Ebene PP-Kontexte erhoben werden, in denen die zielsprachliche Präposition korrekt realisiert wurde, als auch solche, in denen Korrekturen zum Präpositionsgebrauch (z. B. Wahl der Präposition) vorliegen. Mit Abfrage auf Ebene der ZH1 ist es jedoch nicht möglich, diejenigen PP-Kontexte zu erheben, in denen der/die Lernende die Präposition in einem Kontext hinzugefügt hat, in dem zielsprachlich keine PP gefordert ist, d. h. die realisierte Präposition wurde auf Ebene der ZH1 getilgt. Um diese PP-Kontexte ebenfalls bei der Datenauswertung zu berücksichtigen, beziehe ich mich in einer entsprechenden Korpusabfrage auf die ZH1Diff-Ebene und das Editiertag DEL (Näheres s. Kap. 3.3.2). Durch die Kombination aus spezifischer Wortform (in diesem Fall *an* und *auf*²⁴⁰) und dem Editiertag DEL lassen sich die PP-Kontexte inkorrekt Hinzufügungen der ausgewählten Präpositionen im Korpus erheben.

Durch den Bezug auf die Ebene ZH1- bzw. ZH1-Diff werden alle im Korpus enthaltenen Gebrauchskontexte der ausgewählten Präpositionen *an* und *auf* erfasst. Wie in Kapitel 6 beschrieben, liegen in MERLIN Fehlerannotationen im Bereich Präpositionsgebrauch vor. Prinzipiell bestünde somit die Möglichkeit, gezielt entsprechende Fehlerannotationstags abzufragen, um Kontexte zu erheben, in denen die Präpositionen *an* und *auf* inkorrekt realisiert werden. In MERLIN liegen allerdings nicht für alle Lernertexte Fehlerannotationen vor. Mit einer Abfrage der Fehlerannotationskategorien werden folglich nicht alle Gebrauchskontexte der entsprechenden PP im Korpus erfasst, der Gebrauch von Präpositionen und PP in MERLIN könnte folglich nicht in seiner Gänze abgebildet werden. Aus

²⁴⁰ Fortan sind bei Nennung der Präpositionen *an* und *auf* ebenfalls deren Verschmelzungen inbegriffen.

diesem Grund verzichte ich auf eine gezielte Abfrage der Fehlerannotationskategorien. Die Fehlerannotationen, wenn vorliegend, werden jedoch im Rahmen der Datenerhebung aus dem Korpus exportiert.

Die Korpusabfragen zur vorliegenden Fallstudie sind in der folgenden Tabelle (Tab. 44) dargestellt, die Ergebnisse beider Abfragen je Präposition und GeR-Gesamtniveau sind in Tabelle 45²⁴¹ abgebildet. Ich verwende hier den Begriff *PP-Kontexte*, da sowohl zielsprachlich korrekte als auch zielsprachlich nicht korrekte Verwendungen von PP bei der Datenerhebung berücksichtigt werden. Ich vermeide hier die bloße Bezeichnung *PP*, da bei der Datenerhebung (und Datenauswertung) auch solche Treffer berücksichtigt werden, in denen die Lernenden keine PP produziert haben, der ZH1 nach jedoch an entsprechender Stelle eine PP gefordert ist.

Korpusabfrage	Beschreibung
TH1=/[Aa]n [Aa]m [Aa]ns/ TH1=/[Aa]uf [Aa]ufs/ Abfrage je GeR-Gesamtniveau	Abfrage aller Instanzen entsprechender Wortformen in zielsprachlich geforderten Kontexten = Ermittlung von PP-Kontexten, in denen die Präposition auf Ebene der ZH1 nicht getilgt wurde.
tok=/[Aa]n [Aa]m [Aa]ns/_=_TH1Diff="DEL" tok=/[Aa]uf [Aa]ufs/_=_TH1Diff="DEL" Abfrage je GeR-Gesamtniveau	Abfrage aller Instanzen entsprechender Wortformen in zielsprachlich nicht geforderten Kontexten = Ermittlung von PP-Kontexten, in denen die Präposition auf Ebene der ZH1 getilgt wurde.

Tabelle 44: Korpusabfragen zur Fallstudie Syntaktische Funktion der PP.

Präposition (Wortform)	Anzahl Treffer im Korpus									
	GeR-Gesamtniveau									Gesamt
	A1	A2	A2+	B1	B1+	B2	B2+	C1	C2	
<i>an</i>	17	116	84	169	111	218	112	60	3	890
<i>auf</i>	3	37	33	101	64	158	52	34	5	487
Summe	20	153	117	270	175	376	164	94	8	1377

Tabelle 45: Ergebnis der Korpusabfragen je Präposition (Wortform) und GeR-Gesamtniveau (nicht bereinigt).

Die ermittelten 1377 Treffer im Korpus (Tab. 45) werden mit dem im Korpusrecherchesystem ANNIS integrierten GridExporter getrennt nach Präposition und je Präposition getrennt nach GeR-Gesamtniveau exportiert. Je Treffer werden für einen Kontext von zehn Token rechts und links folgende Daten exportiert: Die Original-Lerneräußerung, die entsprechende ZH1 und

²⁴¹ Die Tabelle enthält die Trefferanzahl der jeweils abgefragten Wortform *an/auf* inkl. ihrer Verschmelzungen *am* und *ans* (andere Verschmelzungen wie z. B. *aufs* lagen nicht auf Ebene der ZH1 vor). Die Werte sind nicht bereinigt, die Tabelle enthält u.a. noch Pseudotreffer, die in der Phase der Datenaufbereitung aussortiert werden.

ZH1Diff, die Identifikationsnummer (ID), die Erstsprache²⁴², das GeR-Gesamtniveau sowie, wenn vorliegend, die Fehlerannotation im Bereich Präpositionsgebrauch. Die exportierten Korpusstreffer werden getrennt nach GeR-Gesamtniveau und Präposition in Excel-Dokumente importiert und dort für die weitere Analyse aufbereitet. Auf die Datenaufbereitung gehe ich im Folgenden ein.

9.4 Datenaufbereitung

In diesem Kapitel wird die Datenaufbereitung beschrieben – diese umfasst die Darstellung der Annotationskategorien und die Skizzierung des Annotationsprozesses.

9.4.1 Annotationskategorien

Für die Datenauswertung werden die exportierten Korpusstreffer in Excel-Dokumenten aufbereitet, d. h. nach unterschiedlichen Kategorien annotiert. Im Vordergrund steht die Aussortierung von Pseudotreffern, die Annotationen der syntaktischen Funktionen der im Datensatz vorhandenen PP sowie die Annotationen vorliegender Fehler im Präpositionsgebrauch. Für die Phase der Datenaufbereitung habe ich ein Annotationshandbuch erstellt, in dem die Annotationskategorien und -richtlinien ausführlich beschrieben und anhand von Beispielen illustriert werden. Das finale Annotationshandbuch findet sich im Anhang dieser Arbeit²⁴³. Zur besseren Nachvollziehbarkeit gebe ich die Annotationskategorien zusätzlich in der nachfolgenden Tabelle (Tab. 46) wieder.

Annotationskategorie (Annotationswert 1/0) ²⁴⁴		Erläuterung	Beispiele (z.T. aus MERLIN) ²⁴⁵
Syntaktische Funktion der PP	1	Objekt-PP PP in Objekt-Funktion. Hierzu zähle ich auch Adjektive mit Präpositionsreaktion	<i>Ich denke an Menschen [...]</i> <i>Wenn ich auf diese Frage antworten soll [...]</i>
	2	Adverbiale PP PP in adverbialer Funktion	<i>Man sieht auf die Straße fremde Restaurants, [...]</i> <i>Heutzutage studiere ich an der [...] Hochschule.</i>
	3	Attributive PP PP in attributiver Funktion	<i>Spaß an der Arbeit</i> <i>Antwort auf meine Fragen</i>

²⁴² Für die Datenauswertung wird die Erstsprache der Lernenden nicht weiter berücksichtigt. Dieses Metadatum wird dennoch exportiert, um Informationen über die einzelnen im Datensatz vertretenen Erstsprachen zu erhalten.

²⁴³ Das Annotationshandbuch, so wie im Anhang vorliegend, wurde ebenfalls im Rahmen der Datenaufbereitung zur Fallstudie des Semantischen Gehalts der Präposition eingesetzt (Kap. 10).

²⁴⁴ Der Annotationswert gibt an, welche Werte bei der Annotation vergeben werden konnten; 1 und 0 entsprechen der Zuordnung (1) bzw. Nicht-Zuordnung (0) einer PP in die entsprechende Annotationskategorie (Näheres s. Annotationshandbuch im Anhang).

²⁴⁵ Die Beispiele repräsentieren Original-Lerneräußerungen. Zur besseren Nachvollziehbarkeit ist die ZH1 vereinzelt angegeben.

Weitere Annotationskategorien	4	Rest	PP-Kontexte, deren Zuordnung in andere Kategorien aufgrund unterschiedlicher Gründe nicht möglich war.	<i>Aber am Silvester war schön²⁴⁶</i> <i>ZH1: Aber Silvester war schön.</i>
	5	Pseudotreffer	Form ist keine PP, sondern z. B. eine Verbpartikel oder Bestandteil einer Adjektivform (Superlativ)	<i>Dabei tauchen immer einige Fragen auf, [...]</i> <i>Ich finde es am besten, [...]</i>
	6	Nicht interpretierbar	PP-Kontext (auf Ebene der ZH1) aufgrund von z. B. semantischen Abweichungen nicht interpretierbar bzw. unverständlich ²⁴⁷	<i>Ich fürrt sich auf Sie.</i> <i>ZH1: Ich führe sich auf Sie.</i>
	7	Funktionsverbgefüge²⁴⁸	Prädikatäquivalent aus Funktionsverb und PP	<i>Zur Anklage kommen</i> <i>In Anspruch nehmen</i>
	8	Formelhafte Sequenz	PP, die formelhafte Sequenzen darstellen (s. auch Kap. 1.2.5)	<i>auf Wiedersehen</i> <i>auf jeden Fall</i> <i>auf die Palme bringen</i>

Tabelle 46: Annotationskategorien in der Fallstudie Syntaktische Funktion der PP.

Im Falle vorliegender Objekt-PP werden bei der Annotation zusätzlich die entsprechenden Verben bzw. Adjektive erfasst, das entsprechende Vorgehen ist im Annotationshandbuch beschrieben (s. Anhang).

Zur Beantwortung der Leitfragen dieser Fallstudie sind lediglich PP-Kontexte mit *an/auf* relevant, die als Objekt-PP oder adverbiale PP annotiert werden (s. Tab. 46). Aus diesem Grund werden PP-Kontexte, die bei der Annotation anderen Kategorien zugeordnet werden, von der weiteren Auswertung ausgeschlossen. Zur Bestimmung von Objekt-PP (und zur Abgrenzung dieser von adverbialen PP) werden u.a. linguistische Ressourcen (z. B. E-Valbu oder Duden-Online) konsultiert (Näheres s. Annotationshandbuch im Anhang).

Für die Erfassung vorliegender Fehler und die Annotation der Fehlertypen im Bereich Präpositionsgebrauch definiere ich einen Fehler als eine Abweichung zwischen der Original-

²⁴⁶ Die produzierte PP *am Silvester* wurde auf ZH1 zur NP *Silvester* modifiziert. Diese NP trägt die Funktion des Subjekts im dazugehörigen Satz. Man kann diese Modifikation so interpretieren, dass der Lernende anstelle einer Subjekt-NP (*Silvester*) eine ‚Subjekt-PP‘ (*am Silvester*) produziert hat. Da die syntaktische Funktion *Subjekt* nicht in der vorliegenden Fallstudie behandelt wird, wird solch ein PP-Kontext als *Rest* annotiert. In der Tat lag im gesamten Datensatz der vorliegenden Fallstudie lediglich eine derartige Klassifikation vor. Nähere Informationen zur Annotation von inkorrekt hinzugefügten Präpositionen finden sich im Annotationshandbuch (s. Anhang).

²⁴⁷ Da beim Korpusaufbau auf Ebene der ZH1 lediglich grammatische und orthographische Fehler korrigiert wurden, ist es möglich, dass die ZH1 semantische Abweichungen enthält, die erst auf Ebene der ZH2 korrigiert werden (vgl. MERLIN project 2014). In der vorliegenden Fallstudie beziehe ich mich lediglich auf die Ebene ZH1.

²⁴⁸ Ich möchte hier anmerken, dass im Datensatz der vorliegenden Fallstudie keine PP-Kontexte von FVG vorliegen.

Lerneräußerung und entsprechender ZH (ZH1) im Lernerkorpus. Ein zielsprachlich *inkorrekt*er PP-Kontext liegt demnach vor, wenn auf Ebene der ZH1 im Vergleich zur Original-Lerneräußerung eine Abweichung im Präpositionsgebrauch vorliegt – sei es, dass die betroffene Präposition ergänzt, ersetzt, gelöscht oder verschoben wurde. Als zielsprachlich *korrekter* PP-Kontext gilt in meiner Untersuchung folglich jeder PP-Kontext, in dem auf Ebene der ZH1 im Bereich Präpositionsgebrauch keine Abweichung zur Original-Lerneräußerung vorliegt. Diese Definition bildet in der vorliegenden Fallstudie die Grundlage für die Annotationen und die Auswertungen der Fehlertypen im Bereich Präpositionsgebrauch (s. Kap. 9.5.2).

Für die Datenaufbereitung wird eine weitere Annotationskategorie (Fehlertyp) angesetzt, um die Fehlertypen im Bereich Präpositionsgebrauch zu annotieren. Für die Annotation der Fehlertypen im Bereich Präpositionsgebrauch orientiere ich mich an den Fehlerannotationskategorien in MERLIN. In den einzelnen Excel-Annotationsdokumenten wird, wenn vorliegend, in einer separaten Spalte der Fehlertyp im Gebrauch der Präpositionen *an/auf* erfasst. Die folgende Tabelle (Tab. 47) gibt einen Überblick über die Annotationskategorie *Fehlertyp* und die dazugehörigen Annotationenwerte. Beispiele aus MERLIN sind ebenfalls angegeben.

Annotationskategorie <i>Fehlertyp</i> (Annotationenwert ²⁴⁹)		Beispiel (aus MERLIN)
1	Kein Fehler im Präpositionsgebrauch (0)	[...] weil ich an Ihrer Anzeige interessiert bin. ZH1: [...] weil ich an Ihrer Anzeige interessiert bin.
2	Auslassung einer Präposition (o)	dann warte ich dich. ZH1: Dann warte ich auf dich.
3	Hinzufügung einer Präposition (ad)	[...] suche ich gerne auf eine neue Wohnung. ZH1: [...] ich suche gerne eine neue Wohnung.
4	Wahl der Präposition (ch)	[...] zu den Gastfamilie anpassen. ZH1: [...] an die Gastfamilie anpassen.
5	Position der Präposition (pos)	[...] Informativ Schwerpunkt mit Systemkaufmann. ZH1: [...] Informatik mit Schwerpunkt Systemkaufmann.
6	Zusammenziehung der Präposition (merge)	wie viele Stunden Prowoche [...] ZH1: Wie viele Stunden pro Woche [...]
7	Trennung der Präposition (split)	Da ich zur Zeit keine Familie habe [...] ZH1: Da ich zurzeit keine Familie habe [...]

Tabelle 47: Annotationskategorien und -werte zum Fehlertyp im Bereich Präpositionsgebrauch in der Fallstudie Syntaktische Funktion der PP.

An dieser Stelle sei noch Folgendes bezüglich der Fehlertypen in Tabelle 47 erläuternd hervorgehoben: Für die vorliegende Fallstudie sind vor allem die Fehlertypen 1.-5. von Interesse, da diese die syntaktische Verwendung (bzw. Nicht-Verwendung) der jeweiligen

²⁴⁹ Die Annotationenwerte der Fehlertypen (in Klammern) finden sich ebenfalls in den entsprechenden Fehlerannotationstags in MERLIN wieder (vgl. dazu auch Wisniewski et al. 2014: 12).

Präposition betreffen. Fehlertyp 6 und 7 werden bei der Analyse nicht beachtet, da diese nicht den syntaktischen Gebrauch der Präposition betreffen, sondern lediglich ihre Schreibweise bzw. Anbindung an angrenzende Wortformen. Tatsächlich wurden im MERLIN-Korpus für die Präpositionen *an* und *auf* keine Instanzen der Fehlertypen 6 und 7 ermittelt.

Die Annotationskategorien und -richtlinien zur Annotation der Fehlertypen im Präpositionsgebrauch sind im Annotationshandbuch (s. Anhang) aufgeführt. Diese orientieren sich an den Richtlinien und am Annotationsschema des MERLIN-Korpus. Im Korpus fehlende Fehlerannotationen²⁵⁰ werden manuell ergänzt und für die Datenauswertung berücksichtigt. Im Korpus vorliegende Fehlerannotationen, die als inkorrekt identifiziert werden, werden für die vorliegende Fallstudie korrigiert. Nur die korrigierten Annotationswerte werden für die Datenauswertung berücksichtigt. Da die ZH1 für alle Lernertexte im Korpus vorliegt, ist eine konsistente Fehleridentifikation und -annotation gewährleistet.

9.4.2 Annotationsprozess

Die exportierten PP-Kontexte mit *an* und *auf* wurden gemäß dem Annotationshandbuch von einer studentischen Hilfskraft und mir manuell annotiert. Ein Annotationshandbuch stellt ein wichtiges Instrument in der korpuslinguistischen Arbeit dar, da es die Objektivität des Annotationsprozesses fördert und zur Evaluierung der Annotationskategorien dient (Lemnitzer/Zinsmeister 2015: 101ff.). Eine erste Version des Annotationshandbuchs habe ich in einer Pilotphase (Probeannotationen) getestet und im Anschluss kontinuierlich überarbeitet.

Da Annotationen nur sinnvoll sind, wenn sie konsistent durchgeführt werden (Lüdeling/Hirschmann 2015: 148), ist die Überprüfung der Reliabilität von Annotationen ein wichtiger Schritt jeder korpuslinguistischen Untersuchung. Für die Ermittlung der Reliabilität der Annotationen in der vorliegenden Fallstudie (gemäß den Annotationskategorien aus Tabelle 46) wurde in einer Evaluierungsphase die Übereinstimmung der Annotationen (engl. *Inter-Annotator-Agreement*, kurz *IAA*²⁵¹) ermittelt. Hierfür wurden exemplarisch 100 PP-Kontexte mit der Präposition *an*²⁵² von beiden Annotatorinnen doppelt annotiert. Die Auswahl der 100 PP-Kontexte erfolgte randomisiert und computergestützt. Die ausgewählten PP-Kontexte umfassen die GeR-Gesamtniveaus A2-C1, wobei alle Gesamtniveaus zu gleichen Anteilen repräsentiert sind. Es wurde sichergestellt, dass die 100 PP-Kontexte keine

²⁵⁰ Zur Erinnerung: Von 1033 Lernertexten im MERLIN-Korpus sind lediglich 752 Lernertexte im Bereich Orthographie und Grammatik annotiert, das entspricht ca. 73% aller in MERLIN enthaltenen Texte.

²⁵¹ Auch *inter-coder agreement* (Artstein/Poesio 2008) genannt.

²⁵² Das IAA wurde parallel zu bereits laufenden Annotationen berechnet. Für die Berechnung des IAA wurde lediglich die Präposition *an* gewählt, da PP-Kontexte mit dieser Präposition zur Zeit der IAA-Berechnung noch von keiner der Annotatorinnen annotiert worden waren. Auf diese Weise konnte gewährleistet werden, dass die doppelten Annotationen unbeeinflusst von evtl. bereits bekannten PP-Kontexten waren. PP-Kontexte mit *auf* und der Verschmelzung *am* wurden zum Zeitpunkt der IAA-Berechnung bereits von beiden Annotatorinnen gesichtet, z.T. war die Annotation bereits begonnen worden. In der Fallstudie zum semantischen Gehalt der Präposition (Kap. 10), die im Anschluss an die hier vorliegende Fallstudie durchgeführt wurde, erfolgte die IAA-Berechnung vor der Annotationsphase, sodass dort für das IAA PP-Kontexte mit beiden untersuchten Präpositionen berücksichtigt werden konnten.

Pseudotreffer enthielten. Zur Berechnung des IAA wird das Kappa-Maß (κ) gewählt (vgl. Cohen 1960). Das κ -Maß gilt als Standard zur Evaluierung von Annotationen durch zwei oder mehr AnnotatorInnen (vgl. Lemnitzer/Zinsmeister 2015: 61, Keibel/Perkuhn/Kupietz 2012: 60) bei nominal skalierten Variablen²⁵³. Das κ -Maß berücksichtigt, dass Übereinstimmungen zwischen AnnotatorInnen bzw. Annotationen dem Zufall geschuldet sein können (vgl. auch Lemnitzer/Zinsmeister 2015: 61), zufällige Übereinstimmungen bei den Annotationen werden somit herausgerechnet (vgl. auch Artstein/Poesio 2008: 559ff., Krippendorff 1980: 135). Die Reliabilität der Annotationen spiegelt sich im κ -Wert wider: Ein κ -Wert von 0 entspricht keiner Übereinstimmung, ein κ -Wert von 1 zeigt eine vollständige Übereinstimmung an. Für die 100 PP-Kontexte mit der Präposition *an* wurde in der vorliegenden Fallstudie ein κ -Maß von 0.9 ermittelt²⁵⁴. Carletta (1996: 252) spricht mit Verweis auf Krippendorff (1980: 147) von einem κ -Wert von >0.8 als Repräsentation einer guten Reliabilität. In der vorliegenden Fallstudie zeigt sich mit $\kappa=0.9$ folglich eine hohe Reliabilität der vorliegenden Annotationen.

Im weiteren Annotationsprozess wurden die PP-Kontexte unter der studentischen Hilfskraft und mir aufgeteilt und entsprechend dem Annotationshandbuch in mehreren Phasen annotiert. Nach jeder Annotationsphase wurden im Rahmen persönlicher Gespräche Unklarheiten diskutiert und nach Möglichkeit aufgelöst. Das Annotationshandbuch wurde dabei kontinuierlich überarbeitet und die Annotationen, wenn nötig, überprüft und korrigiert.

Nach Abschluss des Annotationsprozesses wurden die Annotationen nochmals stichprobenartig kontrolliert und im Falle von Abweichungen zu den Annotationsrichtlinien im Annotationshandbuch korrigiert. Die Korrekturen sind jeweils dokumentiert. Die annotierten Forschungsdaten sind auf MADATA²⁵⁵ veröffentlicht.

9.5 Datenauswertung

Im ersten Schritt der Datenauswertung werden die PP-Kontexte ausgeschlossen, die weder als Objekt-PP noch als adverbiale PP annotiert wurden (vgl. Tab. 46). PP-Kontexte der Kategorien Objekt und Adverbiale wurden zunächst dahingehend überprüft, ob sie PP enthalten, die sich ebenfalls in den zugrunde liegenden Aufgabenstellungen²⁵⁶ finden. Ich gehe davon aus, dass PP, die identisch sowohl in den Original-Lerneräußerungen als auch der zugrundeliegenden Aufgabenstellung vorliegen, den Sprachgebrauch des Lernenden nicht adäquat widerspiegeln, da die (plausible) Möglichkeit besteht, dass die Lernenden PP aus den Aufgabenstellungen kopieren und diese in eigenen Produktionen reproduzieren. Um solche PP-Kontexte so gut wie möglich für die Datenauswertung auszuschließen, habe ich

²⁵³ Unter nominal skalierten Variablen versteht man die Zuordnung von Ausprägungen zu Variablen, bei der die Ausprägungen eindeutig einer Variable zugeordnet werden können und keine Rangfolge bei den Ausprägungen vorliegt (vgl. Meindl 2011: 70).

²⁵⁴ Zur Berechnung des κ -Wertes wurde die Website *VassarStats* von Richard Lowry (2001-2009) konsultiert (<http://vassarstats.net/kappa.html>) (Mai 2019).

²⁵⁵ <https://madata.bib.uni-mannheim.de/>

²⁵⁶ Die Aufgabenstellungen sind online verfügbar (MERLIN, https://merlin-platform.eu/C_mcorpus.php#anchor3, Mai 2019).

entsprechende PP für die weitere Datenauswertung manuell aussortiert²⁵⁷. Das Vorgehen möchte ich im Folgenden beschreiben.

Zunächst wurden alle PP mit *an/auf* in den zugrundeliegenden Aufgabenstellungen ermittelt. Im Falle vorliegender Objekt-PP in den Aufgabenstellungen wurden zunächst die Verben bzw. Adjektive erfasst, die eine entsprechende Präpositionsreaktion aufweisen. Im Anschluss wurden die Verben bzw. Adjektive mit entsprechender Präpositionsreaktion mit dem Gebrauch der Verben/Adjektive in den Annotationsdokumenten verglichen; der Gebrauch der Verben/Adjektive wurde im Rahmen der Datenaufbereitung erfasst (s. 9.4.1). Zudem erfolgte eine Kontrolle, ob die Verben/Adjektive aus Texten stammen, die zur entsprechenden Aufgabenstellung (die die zu kontrollierenden Verben/Adjektive enthält) verfasst wurde. Im Falle einer Übereinstimmung wurde der betroffene PP-Kontext für die weitere Datenauswertung aussortiert²⁵⁸. Im Falle adverbialer PP in den Aufgabenstellungen wurde jedes Annotationsdokument nach dem entsprechenden PP-Kontext durchsucht und die entsprechenden PP-Kontexte aussortiert, soweit diese auf Aufgabenstellungen zurückzuführen waren, die die entsprechenden adverbialen PP enthielten. Um Konsistenz bei der Vorgehensweise zu gewährleisten, erfolgte die Durchsuchung der Annotationsdokumente im Falle adverbialer PP-Kontexte stets formbasiert, d. h. ich sortierte jeweils die PP-Treffer aus, die formgleich in den Original-Lerneräußerungen und in den Aufgabenstellungen enthalten waren.

Die folgende Tabelle (Tab. 48) gibt einen Überblick über die PP-Kontexte mit *an* und *auf* je syntaktischer Funktion und GeR-Gesamtniveau, die für die weitere Datenauswertung berücksichtigt werden. Für die weiteren Auswertungen werden bei Angabe des GeR-Gesamtniveaus die Kriterien- und die Plus-Niveaus zusammengefasst (A2 schließt somit A2 und A2+ ein). Die PP-Kontexte der Niveaus A1 sowie C2 werden von der weiteren Auswertung ausgeschlossen, da im Datensatz vergleichsweise wenige Objekt-PP und adverbiale PP mit *an* und *auf* auf diesen GeR-Gesamtniveaus vorliegen²⁵⁹. Insgesamt werden 913 PP-Kontexte von Objekt-PP und adverbialen PP mit *an/auf* ausgewertet.

²⁵⁷ Entsprechende PP-Kontexte habe ich aus pragmatischen Gründen in die Kategorie *Pseudotreffer* einsortiert. Dass es sich um PP-Kontexte aus Aufgabenstellungen handelt, wurde in der Kommentarspalte der entsprechenden Excel-Dokumente vermerkt.

²⁵⁸ An dieser Stelle möchte ich darauf hinweisen, dass das Verb *orientieren (an)* bzw. die entsprechende Objekt-PP mit *an*, die in einer zugrundeliegenden Aufgabenstellung enthalten war, vergleichsweise sehr häufig von den Lernenden verwendet wurde (z.T. innerhalb der Wiedergabe der Aufgabenstellung) und damit einen deutlichen Ausreißer innerhalb des Datensatzes dieser Fallstudie darstellte. Aus pragmatischen Gründen habe ich bei diesem Verb darauf verzichtet, in den einzelnen PP-Kontexten zu diesem Verb in den Annotationsdokumenten zu kontrollieren, ob für jeden PP-Kontext auch tatsächlich die entsprechende Aufgabenstellung zugrunde liegt und habe alle PP-Kontexte des Verbs *orientieren (an)* aus meinem Datensatz entfernt.

²⁵⁹ Die Anzahl entsprechender PP-Kontexte mit *auf* und *an* auf dem Niveau A1 bzw. C2 lag lediglich zwischen eins und zehn PP-Kontexten.

	GeR-Gesamtniveau				
	A2	B1	B2	C1	Summe
an-PP					
Gesamt	148	173	193	29	543
Anzahl Objekt-PP	4	24	76	15	119
Anzahl adverbiale PP	144	149	117	14	424
auf-PP					
Gesamt	62	133	154	21	370
Anzahl Objekt-PP	57	107	87	9	260
Anzahl adverbiale PP	5	26	67	12	110

Tabelle 48: PP-Kontexte je Präposition (an/auf), GeR-Gesamtniveau und syntaktischer Funktion der PP (bereinigt).

Wie Tabelle 48 entnommen werden kann, sind die PP-Kontexte mit den Präpositionen *an* und *auf* ungleich verteilt: Es gibt deutlich mehr PP-Kontexte, die die Präposition *an* enthalten im Vergleich zu PP-Kontexten mit der Präposition *auf*. Innerhalb der jeweiligen PP-Kontexte mit *an* und *auf* sind zudem Objekt-PP und adverbiale PP ebenfalls nicht gleichverteilt. Unter Kapitel 9.2 wurde bereits angemerkt, dass eine Anfrage in der Tübinger Baubank TüBa-D/Z (Version 10.1) zur Verteilung syntaktischer Funktionen von PP zeigt, dass adverbiale PP häufiger belegt sind als Objekt-PP. Somit ist es nicht verwunderlich, dass im vorliegenden Datensatz insgesamt ca. 58% aller PP-Kontexte auf adverbiale PP entfallen und ca. 42% aller PP-Kontexte auf Objekt-PP. Die Abfrage in TüBa-D/Z (Version 10.1) zeigt jedoch auch, dass adverbiale *an*-PP häufiger sind als Objekt-PP mit *an*, das Verhältnis Objekt-PP zu adverbiale PP liegt hier bei ca. 1:1,29. Dahingegen sind Objekt-PP mit *auf* häufiger als adverbiale PP mit *auf*, hier liegt das Verhältnis von Objekt-PP zu adverbiale PP bei 1:0,7²⁶⁰. Diese Verhältnisse werden in ähnlicher Weise im vorliegenden Datensatz aus MERLIN widerspiegelt (s. Tab. 48): Im Datensatz liegen deutlich mehr PP-Kontexte adverbialer PP mit der Präposition *an* vor, während bei der Präposition *auf* deutlich mehr PP-Kontexte von Objekt-PP vertreten sind. Die Verhältnisse beider syntaktischer Funktionen im vorliegenden Datensatz entsprechen zwar nicht genau den entsprechenden Verhältnissen in TüBa-D/Z²⁶¹, es lassen sich jedoch parallele Tendenzen feststellen.

In der vorliegenden Fallstudie wird nicht der Gebrauch dieser spezifischen Präpositionen untersucht, sondern der Einfluss der syntaktischen Funktion von PP (Objekt-PP und adverbiale PP) auf den Gebrauch von Präpositionen. Dieser Einfluss soll beispielhaft für PP-Kontexte mit den Präpositionen *an/auf* analysiert werden. Aus diesem Grund werden in der weiteren Datenauswertung die Präpositionen *an* und *auf* ausschließlich gemeinsam berücksichtigt. In der folgenden Tabelle (Tab. 49) ist die Verteilung der PP-Kontexte je

²⁶⁰ Die entsprechenden Abfragen in TüBa-D/Z (Version 10.1) lieferten folgende Ergebnisse: PP als OPP bzw. FOPP mit *an* lieferte 2822 Treffer; PP als V-MOD mit *an* lieferte 3628 Treffer. Bei PP als OPP bzw. FOPP mit *auf* lagen 4035 Treffer vor während PP als V-MOD mit *auf* 2834 Treffer aufwies.

²⁶¹ Im vorliegenden Datensatz liegt das Verhältnis von Objekt-PP zu adverbialer PP bei 1:3,56 (*an*-PP) bzw. bei 2,36:1 (*auf*-PP), vgl. Tab. 48.

syntaktischer Funktion und GeR-Gesamtniveau erneut angegeben, die beiden Präpositionen sind nun zusammengefasst.

Syntaktische Funktion der PP	GeR-Gesamtniveau				
	A2	B1	B2	C1	Summe
Objekt-PP	61	131	163	24	379
Adverbiale PP	149	175	184	26	534
Gesamt	210	306	347	50	913

Tabelle 49: PP-Kontexte je GeR-Gesamtniveau und syntaktischer Funktion der PP.

Um sicherzustellen, dass im Datensatz keine Ausreißer bei den Lernenden²⁶² vorliegen, wurde zusätzlich die Anzahl unterschiedlicher Lernertexte ermittelt, auf die sich die insgesamt 913 PP-Kontexte insgesamt sowie je GeR-Gesamtniveau und syntaktischer Funktion verteilen. Eine entsprechende Übersicht gibt die nachfolgende Tabelle (Tab. 50).

Syntaktische Funktion der PP	GeR-Gesamtniveau				
	A2	B1	B2	C1	Gesamt
Beide	147	195	197	25	564
Objekt-PP	57	114	122	15	308
Adverbiale PP	113	130	126	18	387

Tabelle 50: Anzahl unterschiedlicher Lernertexte je GeR-Gesamtniveau: Gesamtverteilung und getrennt nach syntaktischer Funktion der PP.

Der obigen Tabelle (Tab. 50) ist zu entnehmen, dass die 913 PP-Kontexte mit *an/auf* insgesamt 564 unterschiedliche Lernertexte, d. h. 564 verschiedene DaF-Lernende repräsentieren. Damit liegt eine Datenbasis vor, die weit über den traditionellen Datensammlungen in der Z/FSE-Forschung liegt, die i.d.R. weit weniger Lernende umfassen (vgl. Granger 2008: 269, Granger et al. 2015: 1). Weiterhin kann Tabelle 50 entnommen werden, dass die im Datensatz enthaltenen 379 Objekt-PP 308 unterschiedliche Lernertexte, d. h. Lernende repräsentieren; die 534 adverbialen PP bilden insgesamt 387 unterschiedliche Lernertexte, d. h. Lernende ab.

Die im finalen Datensatz enthaltenen 913 PP-Kontexte von Objekt-PP und adverbialen PP wurden im Folgenden entsprechend den zentralen Leitfragen ausgewertet. Die Leitfragen seien an dieser Stelle nochmals aufgeführt.

1. Welche Unterschiede zeigen sich bei der Korrektheit im Präpositionsgebrauch innerhalb von PP unterschiedlicher syntaktischer Funktion (Objekt-PP/adverbiale PP)?
2. Welche Fehlertypen im Präpositionsgebrauch treten häufig auf und wie unterscheiden sich diese gemäß der syntaktischen Funktion der PP (Objekt-PP/adverbiale PP)?

Für beide Leitfragen wird jeweils das jeweilige GeR-Gesamtniveau der DaF-Lernenden für die Auswertung berücksichtigt.

²⁶² Unter Ausreißern bei den Lernenden verstehe ich hier, dass die ermittelten PP je GeR-Gesamtniveau sehr wenige Lernertexte, d.h. Lernende repräsentieren.

9.5.1 Analysemethoden

Die ermittelten PP-Kontexte beider syntaktischer Funktionen werden mittels der computerunterstützten Fehleranalyse (computer-aided error analysis, CEA) (Dagneaux/Denness/Granger 1998) sowie der kontrastiven Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (Granger 1996, 2015) ausgewertet. Die nachfolgende Tabelle (Tab. 51) gibt einen Überblick über die in der vorliegenden Fallstudie angewandten Analysemethoden und die damit verbundenen Ziele. Für nähere Informationen zur Anwendung dieser Methoden im Kontext der Dissertation siehe Kapitel 7.1.

Analysemethode	Ziel
Computerunterstützte Fehleranalyse	Ermittlung der Abweichungen (bzw. Korrektheit) im Präpositionsgebrauch (in zielsprachlich geforderten und zielsprachlich nicht geforderten PP-Kontexten) je syntaktischer Funktion der PP (Objekt-PP und adverbiale PP) und GeR-Gesamtniveau.
Kontrastive Interlanguage-Analyse	Kontrastiver Vergleich der Ergebnisse der computerunterstützten Fehleranalyse je GeR-Gesamtniveau und syntaktischer Funktion (Objekt-PP und adverbiale PP).

Tabelle 51: Analysemethoden der Fallstudie zur Variable Syntaktische Funktion der PP: Überblick.

Im folgenden Kapitel werden die Ergebnisse der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse und ihre statistischen Auswertungen präsentiert.

9.5.2 Ergebnisse

In diesem Kapitel werden die Ergebnisse der quantitativen Datenauswertung vorgestellt, die eingangs formulierten Hypothesen überprüft und die entsprechenden Leitfragen beantwortet. Das vorliegende Kapitel ist entsprechend der eingangs formulierten Leitfragen gegliedert: Zuerst gehe ich auf die Korrektheit im Präpositionsgebrauch innerhalb von Objekt-PP und adverbialer PP ein (Leitfrage 1). Im Anschluss daran präsentiere ich die entsprechenden quantitativen Auswertungen der Fehlertypen im Präpositionsgebrauch (Leitfrage 2). Dabei wird jeweils überprüft, ob bzw. inwieweit die syntaktische Funktion der PP einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch ausübt. In der Auswertung zur Korrektheit und den Fehlertypen im Präpositionsgebrauch wird jeweils das GeR-Gesamtniveau der Lernenden berücksichtigt.

9.5.2.1 Auswertung Korrektheit im Präpositionsgebrauch

Die statistische Auswertung der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse liefert in Bezug auf die Korrektheit im Präpositionsgebrauch die folgenden zwei zentralen Ergebnisse:

- 1) Insgesamt kann eine hohe Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch ermittelt werden.
- 2) Es liegt eine Interaktion vor zwischen GeR-Gesamtniveau und syntaktischer Funktion der PP: Beide Variablen beeinflussen sich gegenseitig in ihrem Effekt auf die Korrektheit im

Präpositionsgebrauch. Die Korrektheit im Präpositionsgebrauch lässt sich somit nur teilweise auf den Einfluss der syntaktischen Funktion zurückführen. Ein eindeutiger Effekt der syntaktischen Funktion auf die Korrektheit lässt sich nur im Niveau elementarer Sprachverwendung (A2) ermitteln.

Die folgende Tabelle (Tab. 52) fasst zunächst die in der Datengrundlage ermittelten Werte zur Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP mit *an/auf* in einem Überblick zusammen.

Syntaktische Funktion	PP-Kontexte (Gesamt)		Inkorrekte PP-Kontexte		Korrekte PP-Kontexte	
	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent
Objekt-PP	379	100%	123	32%	256	68%
Adverbiale PP	534	100%	107	20%	427	80%

Tabelle 52: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP: Überblick.

Die Werte in Tabelle 52 zeigen, dass die meisten PP-Kontexte mit *an/auf* in Bezug auf den Präpositionsgebrauch unabhängig von der syntaktischen Funktion eine korrekte Verwendung aufweisen. Insgesamt sind 683 PP der 913 PP korrekt realisiert, dies entspricht einem Anteil von ca. 75%. Da in dieser Dissertation die Schwierigkeiten im Präpositionsgebrauch das Forschungsanliegen darstellen, stehen hier jedoch vor allem diejenigen PP-Kontexte im Vordergrund, in denen ein inkorrekt gebrauchter Präpositionen ermittelt werden kann. Der Anteil inkorrekt gebrauchter Präpositionen nimmt immerhin insgesamt ein Viertel (25%) innerhalb der hier untersuchten PP-Kontexte ein. Wie in der obigen Tabelle (Tab. 52) jedoch deutlich wird, liegt die Korrektheit im Präpositionsgebrauch in PP-Kontexten adverbialer PP mit *an/auf* insgesamt deutlich höher (80%) als in PP-Kontexten von Objekt-PP mit *an/auf* (68%).

Die nachfolgende Tabelle (Tab. 53) gibt einen Überblick zu korrekten und inkorrekten PP-Kontexten je syntaktischer Funktion und GeR-Gesamtniveau. Die Werte geben die Verteilung im vorliegenden Datensatz wieder. Die Anzahl unterschiedlicher Lernertexte, d. h. Lernender ist ebenfalls jeweils angegeben.

GeR-Gesamtniveau	Syntaktische Funktion der PP					
	Objekt-PP (O)			Adverbiale PP (A)		
	Korrekte PP-Kontexte (%)	Inkorrekte PP-Kontexte (%)	AL	Korrekte PP-Kontexte (%)	Inkorrekte PP-Kontexte (%)	AL
A2	21 (34%)	40 (66%)	57	114 (77%)	35 (23%)	113
B1	91 (69%)	40 (31%)	114	140 (80%)	35 (20%)	130
B2	123 (75%)	40 (25%)	122	149 (81%)	35 (19%)	126
C1	21 (87%)	3 (13%)	15	24 (92%)	2 (8%)	18
Summe	256 (68%)	123 (32%)	308	427 (80%)	107 (20%)	387

Tabelle 53: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau. AL=Anzahl Lernertexte.

Die Werte in Tabelle 53 zeigen, dass Objekt-PP eine geringere Korrektheit im Bereich Präpositionsgebrauch aufzeigen als adverbiale PP, und das auf jedem hier untersuchten GeR-Gesamtniveau. Zudem wird deutlich, dass der Anteil korrekter PP-Kontexte mit Zunahme des Kompetenzniveaus (von A2 bis C1) für beide syntaktischen Funktionen stetig zunimmt (von 34% auf 87% bzw. von 77% auf 92%).

Die vorliegenden Daten werden statistisch ausgewertet. Hierfür wähle ich eine Regressionsanalyse und berechne gemischte Modelle (Näheres s. Kap. 7.2). Als Zielvariable (abhängige Variable) wird die Korrektheit im Präpositionsgebrauch festgelegt, die syntaktische Funktion der PP (PP_Funktion) sowie das GeR-Gesamtniveau der Lernenden (Kompetenzniveau.op) werden als unabhängige Variablen berücksichtigt und stellen die festen Faktoren dar. Der/die Lernende wird als Zufallsfaktor festgelegt (1|Lerner). Auf diese Weise kann der individuelle Einfluss einzelner Lernender berücksichtigt und herausgerechnet werden. Um zu bestimmen, ob die syntaktische Funktion und das GeR-Gesamtniveau sich in Bezug auf die Korrektheit gegenseitig beeinflussen, wird in das erste Regressionsmodell (RM1) eine Interaktion zwischen diesen beiden kategorialen Variablen (syntaktische Funktion und GeR-Gesamtniveau) integriert. Das entsprechende Regressionsmodell (RM1) ist im Folgenden angegeben:

RM1:

```
glmer(Korrektheit ~ PP_Funktion*Kompetenzniveau.op + (1|Lerner), data = dat, family = "binomial", control = glmerControl(optimizer = "bobyqa"))
```

Die statistische Auswertung zeigt, dass in der Tat eine allgemeine Interaktion zwischen der syntaktischen Funktion der PP und dem GeR-Gesamtniveau vorliegt ($p=0.002595$). Die Visualisierung des Modell-Outputs von RM1 findet sich in Abbildung 38. Aus der Y-Achse ist die Wahrscheinlichkeit für die Zielvariable abgebildet (hier: Korrektheit im Präpositionsgebrauch), auf der X-Achse sind die einzelnen GeR-Gesamtniveaus angegeben. Die durchgezogene Linie repräsentiert die ermittelten Wahrscheinlichkeiten für die Korrektheit im Präpositionsgebrauch für adverbiale PP, die gestrichelte Linie zeigt die entsprechenden Wahrscheinlichkeiten für Objekt-PP. Für jedes GeR-Gesamtniveau ist der Mittelwert zur Wahrscheinlichkeit der Korrektheit angegeben (Kreis) sowie die Standardabweichung zum jeweiligen Mittelwert (senkrechte Linie mit waagerechten Linien).

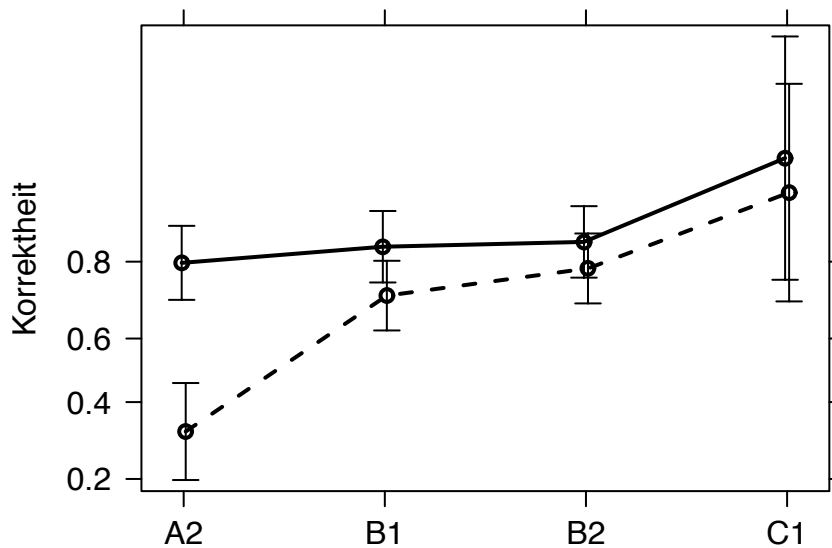


Abbildung 38: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM1. Objekt-PP = gestrichelte Linie, adverbiale PP = durchgezogene Linie.

Abbildung 38 ist zu entnehmen, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch für adverbiale PP konstant höher liegt als für Objekt-PP. Das bedeutet, dass die Präpositionen *an/auf* in Objekt-PP konstant mehr Unsicherheit, d. h. mehr Fehler im Präpositionsgebrauch bei DaF-Lernenden abbilden als dieselben Präpositionen in adverbialen PP. Gleichzeitig wird in Abbildung 38 deutlich, dass die ermittelten Wahrscheinlichkeiten für Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP unterschiedlich stark divergieren und die Wahrscheinlichkeiten für Korrektheit im Präpositionsgebrauch in Objekt-PP und adverbialen PP sich mit zunehmenden GeR-Gesamtniveau deutlich annähern.

Des Weiteren lässt sich für beide syntaktischen Funktionen von PP beobachten, dass die Wahrscheinlichkeiten für Korrektheit im Präpositionsgebrauch mit steigendem GeR-Gesamtniveau stetig zunehmen: Im GeR-Gesamtniveau A2 liegt die Wahrscheinlichkeit für Korrektheit für Objekt-PP bei ca. 30% während sie für adverbiale PP ca. 80% beträgt (Abb. 38). Dahingegen liegen z. B. die Wahrscheinlichkeiten auf dem GeR-Gesamtniveau B2 nah beieinander (knapp über 80% für adverbiale PP und knapp unter 80% für Objekt-PP). Dies deutet darauf hin, dass die syntaktische Funktion der PP insgesamt einen unterschiedlichen Einfluss auf die Korrektheit ausübt, und zwar in Abhängigkeit vom GeR-Gesamtniveau. Die Berechnung der allgemeinen Interaktion, d. h. dem Zusammenwirken von syntaktischer Funktion der PP und GeR-Gesamtniveau, bestätigt dies auch (s.o.): Es liegt eine signifikante allgemeine Interaktion vor zwischen der syntaktischen Funktion der PP und dem GeR-Gesamtniveau. Das bedeutet, dass die Korrektheit im Präpositionsgebrauch durch das Zusammenspiel beider hier untersuchten Variablen prognostiziert werden kann und die syntaktische Funktion der PP je GeR-Gesamtniveau einen unterschiedlichen Einfluss auf die Korrektheit im Präpositionsgebrauch hat. Wenn die syntaktische Funktion der PP in jedem GeR-Gesamtniveau einen ähnlichen oder keinen Effekt auf die Korrektheit im Präpositionsgebrauch hätte, würde keine Interaktion zwischen den Variablen vorliegen.

Die Evaluation des Regressionsmodells RM1 liefert einen Wert von $R^2=0.1690051$. Das bedeutet, dass das Regressionsmodell (RM1) die Korrektheit im Präpositionsgebrauch so wie im Datensatz vorliegend zu ca. 17% erklären kann. Dieser Wert erscheint zunächst nicht besonders hoch, er muss jedoch im Kontext der vorliegenden Fallstudie interpretiert werden. Die Fallstudie untersucht, ob bzw. inwieweit sich Gebrauchsunterschiede in der Verwendung der Präpositionen *an/auf* in Objekt-PP und adverbialen PP feststellen lassen und ob hierbei ein Effekt der syntaktischen Funktion und des GeR-Gesamtniveaus festgestellt werden kann. Es geht nicht darum, anzunehmen, dass die Korrektheit im Gebrauch von *an/auf* ausschließlich auf den Einfluss der syntaktischen Funktion und des GeR-Gesamtniveaus zurückzuführen ist. Es soll jedoch ermittelt werden, ob überhaupt ein Effekt der syntaktischen Funktion und des GeR-Gesamtniveaus festgestellt werden kann. Die Varianz in lernersprachlichen Produktionen wird durch eine Vielzahl unterschiedlicher Faktoren beeinflusst (vgl. Ellis 1994: 49). Die Berechnungen zum vorliegenden Regressionsmodell (RM1) zeigen demnach, dass 17% der Varianz in der Korrektheit des Präpositionsgebrauchs im vorliegenden Datensatz mithilfe des Regressionsmodells, d. h. allein auf Basis der Variablen *Syntaktische Funktion der PP* und *GeR-Gesamtniveau*, erklärt werden kann.

Der Einfluss der syntaktischen Funktion der PP auf die Korrektheit wird in der Abbildung des Modell-Outputs (Abbildung 38) vor allem auf dem Niveau A2 deutlich – hier unterscheiden sich die ermittelten Wahrscheinlichkeiten je syntaktischer Funktion sehr stark voneinander: Für adverbiale PP liegt eine höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch vor (ca. 80%), bei Objekt-PP ist die Wahrscheinlichkeit für korrekten Präpositionsgebrauch vergleichsweise gering (ca. 30%). Der deutliche Unterschied auf dem GeR-Gesamtniveau A2 kann ebenfalls anhand der jeweiligen Markierungen der Standardabweichungen abgelesen werden – die Markierungen liegen weit auseinander und überlappen nicht den jeweiligen Mittelwert. Im Gegensatz dazu liegen die prognostizierten Wahrscheinlichkeiten für korrekten Präpositionsgebrauch auf den Kompetenzniveaus B2 und C1 sehr nah beieinander – hier überlappen die Markierungen der Standardabweichungen jeweils die entsprechenden Mittelwerte. Dennoch weisen adverbiale PP auf den Niveaus B2 und C1 tendenziell eine etwas höhere Wahrscheinlichkeit für Korrektheit auf als Objekt-PP. Das GeR-Gesamtniveau B1 liegt zwischen diesen Tendenzen und zeigt einen Unterschied zwischen Objekt-PP und adverbialen PP an, der zwar kleiner ist als der entsprechende Unterschied auf dem GeR-Gesamtniveau A2, jedoch etwas größer als der entsprechende Unterschied auf den GeR-Gesamtniveaus B2 und C1. Dies wird ebenfalls an den Markierungen der Standardabweichungen deutlich, die die jeweiligen Mittelwerte nur knapp nicht überlappen. Ein deutlicher Effekt der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch lässt sich den Ergebnissen nach folglich lediglich für das Niveau der elementaren Sprachverwendung (A2) ermitteln. Auf den restlichen GeR-Gesamtniveaus lässt sich (nahezu) kein eindeutiger Effekt der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch feststellen. Lernende zeigen hier für beide syntaktischen Funktionen der PP eine ähnliche Wahrscheinlichkeit für die Korrektheit im Präpositionsgebrauch. Die syntaktische Funktion der PP hat hier folglich nur einen bedingten Einfluss.

Zusammenfassung

Die erste Leitfrage kann den Ergebnissen nach wie folgt beantwortet werden:

1. Welche Unterschiede zeigen sich bei der Korrektheit im Präpositionsgebrauch innerhalb von PP unterschiedlicher syntaktischer Funktion (Objekt-PP/adverbiale PP)?

Im Vergleich von Objekt-PP und adverbialer PP zeigt sich ein Unterscheid hinsichtlich der Korrektheit im Präpositionsgebrauch von PP-Kontexten mit den ausgewählten Präpositionen *an/auf*. Die statistischen Auswertungen zeigen, dass adverbiale PP mit *an/auf* eine tendenziell höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch zeigen als Objekt-PP mit *an/auf*. Die ermittelten Wahrscheinlichkeiten für Objekt-PP und adverbiale PP unterscheiden sich jedoch unterschiedlich stark je GeR-Gesamtniveau der Lernenden. Mittels der statistischen Berechnungen lässt sich eine Interaktion zwischen syntaktischer Funktion und GeR-Gesamtniveau ermitteln. Ein starker Einfluss der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch liegt auf dem GeR-Gesamtniveau A2 vor und zwar dergestalt, dass Objekt-PP hier eine deutlich geringere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch aufweisen. Ein entsprechender, jedoch geringerer Effekt kann auf dem GeR-Gesamtniveau B1 beobachtet werden. Auf den GeR-Gesamtniveaus B2 und C1 dahingegen weisen adverbiale PP und Objekt-PP mit *an/auf* eine nahezu identische Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch auf.

Damit werden die anfangs aufgestellten Hypothesen (1./2.) den Ergebnissen nach nur z.T. bestätigt: Zwar lässt sich tendenziell eine geringere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch bei Objekt-PP ermitteln. Ein eindeutiger Effekt der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch bei DaF-Lernenden liegt jedoch lediglich für das Niveau der elementaren Sprachverwendung (A2) vor. Auf den anderen GeR-Gesamtniveaus liegen die Wahrscheinlichkeiten für Korrektheit im Präpositionsgebrauch bei Objekt-PP und adverbialen PP sehr nah beieinander, hier lässt sich kein entsprechender Effekt feststellen. Die Auswertung bestätigt jedoch die 3. Hypothese, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch mit zunehmendem GeR-Gesamtniveau ansteigt, und zwar unabhängig von der syntaktischen Funktion der PP.

9.5.2.2 Auswertung Fehlertypen im Präpositionsgebrauch

Für die Auswertung der Fehlertypen im Präpositionsgebrauch werden ausschließlich die inkorrekten PP-Kontexte mit *an/auf* (Objekt-PP und adverbiale PP) berücksichtigt. Die statistischen Auswertungen der computerunterstützten Fehleranalyse sowie der kontrastiven Interlanguage-Analyse liefern in Bezug auf die Fehlertypen im Präpositionsgebrauch folgende zentrale Ergebnisse:

- 1) DaF-Lernende zeigen zwei dominante Fehlertypen im Präpositionsgebrauch: Die Auslassung von Präpositionen und die Wahl von Präpositionen.
- 2) In Objekt-PP ist die Wahrscheinlichkeit für Auslassungen signifikant höher als in adverbialen PP und die Wahrscheinlichkeit für eine inkorrekte Präpositionswahl signifikant geringer. Dieses Ergebnis gilt unabhängig vom GeR-Gesamtniveau der Lernenden.

- 3) Mit steigendem GeR-Gesamtniveau nimmt die Wahrscheinlichkeit für den Fehlertyp Auslassung ab; die Wahrscheinlichkeit für den Fehlertyp Wahl nimmt dahingegen zu. Dieses Ergebnis gilt unabhängig von der syntaktischen Funktion der PP.

In Tabelle 53 habe ich bereits die Verteilung der inkorrekt gebrauchten PP-Kontexte je syntaktischer Funktion und GeR-Gesamtniveau präsentiert. Die folgende Tabelle (Tab. 54) gibt die Anzahl inkorrekt gebrauchter PP-Kontexte je syntaktischer Funktion der PP und GeR-Gesamtniveau erneut wieder. Die Tabelle enthält ebenfalls Angaben zur Anzahl unterschiedlicher Lernertexte, d. h. Lernender, die im Datensatz inkorrekt gebrauchter PP-Kontexte repräsentiert sind.

GeR-Gesamtniveau	Syntaktische Funktion der PP		AL Gesamt
	Objekt-PP	Adverbiale PP	
	Inkorrekte PP-Kontexte	Inkorrekte PP-Kontexte	
A2	40	35	64
B1	40	35	66
B2	40	35	64
C1	3	2	3
Summe	123	107	197

Tabelle 54: Inkorrekte PP-Kontexte je syntaktischer Funktion der PP und GeR-Gesamtniveau: Überblick. AL=Anzahl Lernertexte.

Die Werte der unterschiedlichen Lernertexte bzw. Lernenden in Tabelle 54 zeigen, dass die fehlerhaften PP-Kontexte nicht von einigen wenigen DaF-Lernenden produziert wurden, sondern dass vielen unterschiedlichen Lernenden Fehler im Präpositionsgebrauch unterlaufen sind. Im Datensatz der inkorrekten PP-Kontexte sind insgesamt 197 DaF-Lernende repräsentiert (s. Tab. 54).

Für die Auswertung der inkorrekten PP-Kontexte mit *an/auf* werden die im MERLIN vorliegenden (korrigierten) und manuell hinzugefügten Fehlerannotationen im Bereich Präpositionsgebrauch ausgewertet (s. Kap. 9.4.1). Die folgende Tabelle (Tab. 55) gibt zunächst einen Überblick über die absolute und relative Verteilung dieser Fehlertypen je syntaktischer Funktion der PP-Kontexte mit *an/auf* für die GeR-Gesamtniveaus A2, B1, B2 und C1.

Syntaktische Funktion der PP	Fehlertyp im Präpositionsgebrauch				
	Wahl	Auslassung	Position	Hinzufügung	Summe
Objekt-PP	40 (33%)	63 (51%)	1 (1%)	19 (15%)	123 (100%)
Adverbiale PP	63 (59%)	25 (23%)	3 (3%)	16 (15%)	107 (100%)

Tabelle 55: Fehlertypen im Präpositionsgebrauch je syntaktischer Funktion (GeR-Gesamtniveaus A2, B1, B2 und C1): Überblick.

Tabelle 55 lässt sich entnehmen, dass die Fehlertypen Wahl und Auslassung zu den häufigsten Fehlertypen im Präpositionsgebrauch zählen. Insgesamt machen diese Fehlertypen ca. 83% (191 von 230) aller Präpositionsfehler im vorliegenden Datensatz aus. Die Fehlertypen Wahl und Auslassung dominieren auch bei beiden syntaktischen Funktionen der PP. Das bedeutet,

dass DaF-Lernende Präpositionen vorrangig entweder auslassen oder zielsprachlich inkorrekt wählen, vgl. Beispiele 1./2. (Auslassung) bzw. 3./4. (Wahl). Präpositionen werden dahingegen äußerst selten inkorrekt positioniert (Beispiel 5.), jedoch anteilig relativ häufig in syntaktischen Kontexten hinzugefügt, in denen sie zielsprachlich nicht gefordert sind (Beispiel 6.). Die Beispiele stammen aus dem Datensatz der vorliegenden Fallstudie.

1. *Deshalb muss ich mein Kind aufpassen.* (ID: 1061_0120456)
ZH1: Deshalb muss ich auf mein Kind aufpassen.
2. *In Freizeit möchte ich viele Aktivitäten teilnehmen* (ID: 1023_0109914)
ZH1: In der Freizeit möchte ich an vielen Aktivitäten teilnehmen
3. *[...], aber ich habe nichts über diese Thema in ihren Website gefunden* (ID: 1023_0109588)
ZH1: [...], aber ich habe nichts über dieses Thema auf Ihrer Website gefunden.
4. *Bei dem ersten Tag, [...]* (ID: 1023_0108751)
ZH1: An dem ersten Tag, [...]
5. *[...] diese Gedichte von Goethe [...] auf andere Englisch oder Spanisch wiedergeben?* (ID: 1031_0003310)
ZH1: [...] diese Gedichte von Goethe [...] anderen auf Englisch oder Spanisch wiedergeben?
6. *[...] leichter, auf Problemen beim Wohen oder Haussuche zu vermeiden* (ID: 1031_0003354)
ZH1: [...] leichter, Probleme beim Wohnen oder der Haussuche zu vermeiden

Da die letztgenannten Fehlertypen (Position und Hinzufügung) im vorliegenden Datensatz nicht dominant sind, beschränke ich mich in den folgenden Auswertungen ausschließlich auf die dominanten Fehlertypen Wahl und Auslassung. Die untenstehende Tabelle (Tab. 56) gibt die Verteilung der Fehlertypen Wahl und Auslassung je syntaktischer Funktion und GeR-Gesamtniveau wie im Datensatz vorliegend wieder.

GeR-Gesamtniveau	Syntaktische Funktion der PP			
	Objekt-PP		Adverbiale PP	
	Wahl	Auslassung	Wahl	Auslassung
A2	8	28	15	16
B1	9	25	19	7
B2	22	10	27	2
C1	1	0	2	0
Summe	40	63	63	25

Tabelle 56: Fehlertypen Wahl und Auslassung im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau.

Wie man Tabelle 56 entnehmen kann, liegen auf dem GeR-Gesamtniveau C1 keine einzige Instanz des Fehlertyps Auslassung bzw. lediglich drei Instanzen des Fehlertyps Wahl vor. Aufgrund dieser sehr niedrigen Werte wird das GeR-Gesamtniveau C1 für die weitere Auswertung der Fehlertypen im Präpositionsgebrauch ausgeschlossen. Somit werden insgesamt 120 PP-Kontexte von Objekt-PP bzw. 105 PP-Kontexte adverbialer PP von insgesamt 194 verschiedenen DaF-Lernenden ausgewertet (vgl. Tab. 54). Im Fokus stehen die Fehlertypen Wahl und Auslassung bzw. der Einfluss der syntaktischen Funktion und des GeR-Gesamtniveaus auf diese Fehlertypen im Präpositionsgebrauch.

Die inkorrekten PP-Kontexte werden statistisch ausgewertet. Für die statistische Auswertung nutze ich Regressionsanalysen und berechne je Fehlertyp gemischte Modelle (Näheres s. Kap. 7.2). Als Zielvariable (abhängige Variable) wird dabei jeweils der Fehlertyp (Auslassung bzw. Wahl) festgelegt. Als feste Faktoren werden die (unabhängigen) kategorialen Variablen *Syntaktische Funktion der PP* (PP_Funktion) und das *GeR-Gesamtniveau* (Kompetenzniveau.op) gewählt. Wie auch im Regressionsmodell zur Korrektheit im Präpositionsgebrauch wird in den Regressionsmodellen zu den Fehlertypen Auslassung und Wahl der/die Lernende als Zufallsfaktor („random effect“) integriert (1|Lerner), d. h. der Einfluss einzelner Lernender wird berücksichtigt bzw. für die Modelle „herausgerechnet“. Mittels der einzelnen Regressionsanalysen wird überprüft, ob auf Basis des vorliegenden Datensatzes ein Effekt der syntaktischen Funktion der PP und/oder des GeR-Gesamtniveaus auf den Fehlertyp im Präpositionsgebrauch ermittelt werden kann. Im Modell werden ausschließlich inkorrekte PP-Kontexte berücksichtigt (in den Regressionsmodellen sind diese jeweils mit der Referenz *dat.subset* repräsentiert).

Im Folgenden sind die Regressionsmodelle für den Fehlertyp Auslassung (RM2) und für den Fehlertyp Wahl (RM3) und ihre jeweiligen Outputs angegeben. In diesen Regressionsmodellen wird zunächst jeweils überprüft, ob eine Interaktion zwischen den Variablen *Syntaktischer Funktion der PP* und *GeR-Gesamtniveau* vorliegt.

RM2:

Auslassung ~ PP_Funktion*Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.1926	0.3681	-0.523	0.60083
PP_FunktionO	1.0955	0.5399	2.029	0.04246
Kompetenzniveau.opB1	-1.2736	0.5871	-2.169	0.03007
Kompetenzniveau.opB2	-2.7306	0.8513	-3.207	0.00134
PP_FunktionO:Kompetenzniveau.opB1	0.9100	0.7631	1.193	0.23306
PP_FunktionO:Kompetenzniveau.opB2	0.6480	0.9820	0.660	0.50935

RM3:

Wahl ~ PP_Funktion*Kompetenzniveau.op + (1|Lerner), data = dat.subset²⁶³, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.29803	0.36572	-0.815	0.41512
PP_FunktionO	-1.15519	0.56277	-2.053	0.04010
Kompetenzniveau.opB1	0.47739	0.51462	0.928	0.35358
Kompetenzniveau.opB2	1.57071	0.58148	2.701	0.00691
PP_FunktionO:Kompetenzniveau.opB1	-0.33026	0.76513	-0.432	0.66601
PP_FunktionO:Kompetenzniveau.opB2	0.09375	0.76801	0.122	0.90285

Weitere statistische Berechnungen zeigen, dass weder für den Fehlertyp Auslassung (RM2, $p=0.4741$) noch für den Fehlertyp Wahl (RM3, $p=0.8425233$) eine allgemeine Interaktion zwischen den Variablen *Syntaktische Funktion der PP* und *GeR-Gesamtniveau* ermittelt werden kann.

Die Ergebnisse der Regressionsmodelle RM2 und RM3 sind in der folgenden Abbildung (Abb. 39) visualisiert. Zur Abbildung: Auf der Y-Achse ist jeweils die Wahrscheinlichkeit für die Zielvariable angegeben (hier: Fehlertyp Wahl bzw. Fehlertyp Auslassung), auf der X-Achse findet man die einzelnen GeR-Gesamtniveaus. Die durchgezogene Linie bildet jeweils die Wahrscheinlichkeiten für den entsprechenden Fehlertyp für Kontexte adverbialer PP mit *an/auf* ab, die gestrichelte Linie repräsentiert Wahrscheinlichkeiten für den entsprechenden Fehlertyp für Kontexte von Objekt-PP mit *an/auf*.

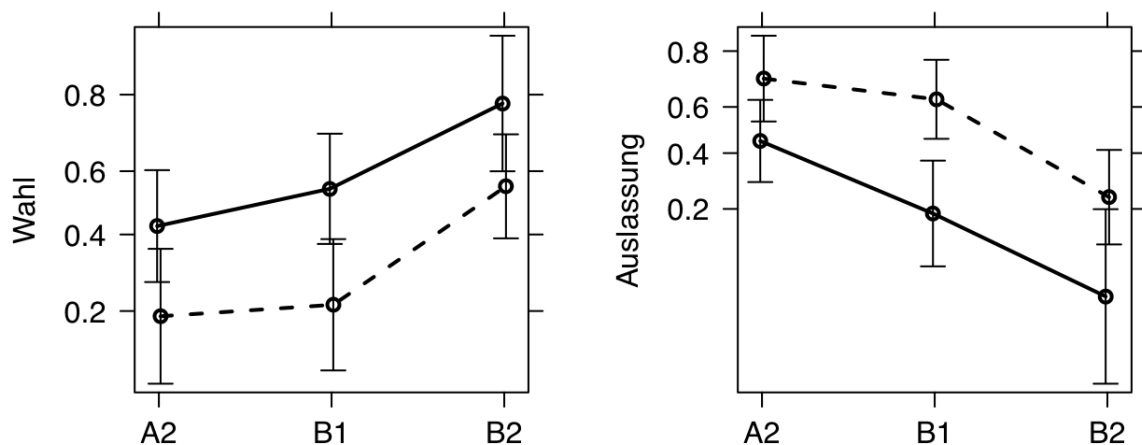


Abbildung 39: Fehlertypen Wahl (links) und Auslassung (rechts) im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau: Ergebnisse der Regressionsmodelle RM2 und RM3. Objekt-PP = gestrichelte Linie, adverbialer PP = durchgezogene Linie.

²⁶³ Die Angabe *dat.subset* referiert auf eine Teilmenge des Datensatzes – die fehlerhaften PP-Kontexte – die im Regressionsmodell berücksichtigt wird.

Die fehlende Interaktion zwischen den Variablen *Syntaktische Funktion der PP* und *GeR-Gesamtniveau* wird ebenfalls in den Visualisierungen der entsprechenden Regressionsmodelle deutlich: In jedem GeR-Gesamtniveau (hier: A2 bis B2) lässt sich feststellen, dass die Unterschiede in den Wahrscheinlichkeiten für den jeweiligen Fehlertyp je syntaktischer Funktion nahezu konstant sind – die entsprechenden Linien für A und O verlaufen je Fehlertyp nahezu parallel, in ähnlich großen Abständen.

Die Visualisierungen der Regressionsmodelle RM2 und RM3 (Abb. 39) zeigen deutlich, dass adverbiale PP (A) eine konstant höhere Wahrscheinlichkeit für den Fehlertyp Wahl aufweisen als Objekt-PP (O) während für letztere eine konstant höhere Wahrscheinlichkeit für Auslassungen und eine konstant niedrigere Wahrscheinlichkeit für wahlbedingte Präpositionsfehler ermittelt werden kann. Die entsprechenden Markierungen der Standardabweichungen überlappen in keinem Falle die entsprechenden Mittelwerte. Dies spiegelt in jedem GeR-Gesamtniveau einen signifikanten Effekt der syntaktischen Funktion auf die Wahrscheinlichkeiten des jeweiligen Fehlertyps wider.

Da in den Regressionsmodellen RM2 und RM3 keine Interaktion zwischen den Variablen *Syntaktische Funktion der PP* und *GeR-Gesamtniveau* ermittelt werden kann, werden die Regressionsmodelle entsprechend modifiziert und die Interaktion jeweils aus den Regressionsmodellen entfernt. In den neuen Regressionsmodellen (RM2a und RM3a) werden die syntaktische Funktion der PP und das GeR-Gesamtniveaus nun lediglich als Einzeleffekte berücksichtigt. Die Modelle sind im Folgenden mit dem jeweiligen Output angegeben:

RM2a:

Auslassung ~ PP_Funktion + Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.4301	0.3236	-1.329	0.1837
PP_FunktionO	1.5812	0.3927	4.027	5.65e-05
Kompetenzniveau.opB1	-0.7826	0.3913	-2.000	0.0455
Kompetenzniveau.opB2	-2.3830	0.5267	-4.524	6.06e-06

RM3a:

Wahl ~ PP_Funktion + Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.2608	0.3059	-0.853	0.393930
PP_FunktionO	-1.2363	0.3426	-3.608	0.000308
Kompetenzniveau.opB1	0.3271	0.3797	0.861	0.389030
Kompetenzniveau.opB2	1.6393	0.4232	3.874	0.000107

Die Visualisierungen beider modifizierter Regressionsmodelle (RM2a und RM3a) sind in Abbildungen 40 und 41 dargestellt. Sie bilden jeweils den Einzeleffekt der syntaktischen Funktion (links) und des GeR-Gesamtniveaus (rechts) auf den Fehlertyp Auslassung (Abb. 40) bzw. Fehlertyp Wahl (Abb. 41) ab. Beide Abbildungen machen nochmals die Einzeleffekte, die bereits in Abbildung 39 ableitbar waren, deutlich sichtbar: Der Fehlertyp Auslassung ist signifikant wahrscheinlicher in Objekt-PP zu beobachten (ca. 50% zu ca. 20%) während der Fehlertyp Wahl signifikant wahrscheinlicher in adverbialen PP (60% zu 30%) ist (Abb. 40 und 41). Dieses Ergebnis gilt unabhängig vom GeR-Gesamtniveau. Weiterhin zeigen die Abbildungen 40 und 41, dass die Wahrscheinlichkeit für den Fehlertyp Auslassung mit steigendem Kompetenzniveau sinken (Abb. 40). Im Gegensatz dazu steigt die Wahrscheinlichkeit für den Fehlertyp Wahl (Abb. 41). Diese Entwicklungen der Wahrscheinlichkeiten einzelner Fehlertypen im Kompetenzverlauf gelten jeweils für beide syntaktischen Funktionen und sind auf den Effekt des GeR-Gesamtniveaus zurückführbar. Für beide Fehlertypen kann jeweils ein allgemeiner signifikanter Effekt des GeR-Gesamtniveaus ermittelt werden (für Wahl $p=0.0002744$, für Auslassung $p=3.496e-05$).

Es wird somit deutlich, dass die Wahrscheinlichkeiten für die Fehlertypen Auslassung und Wahl im Gebrauch von *an/auf* sowohl durch die syntaktische Funktion der PP (Objekt-PP bzw. adverbiale PP) als auch durch das GeR-Gesamtniveau prognostiziert werden können; In beiden Regressionsmodellen (RM2a und RM3a) können entsprechende Einzeleffekte ermittelt werden.

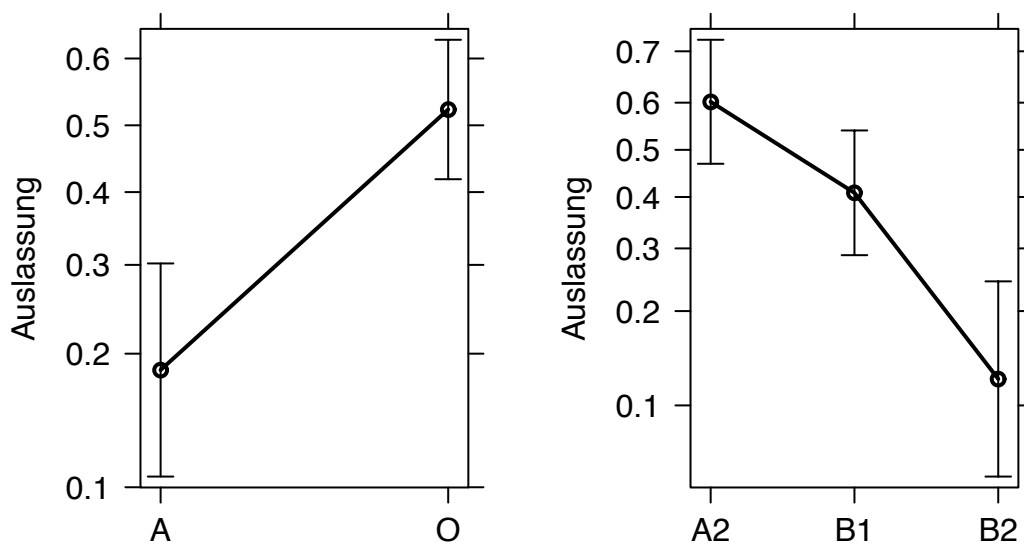


Abbildung 40: Fehlertyp Auslassung je syntaktischer Funktion der PP und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM2a. A = adverbiale PP, O = Objekt-PP.

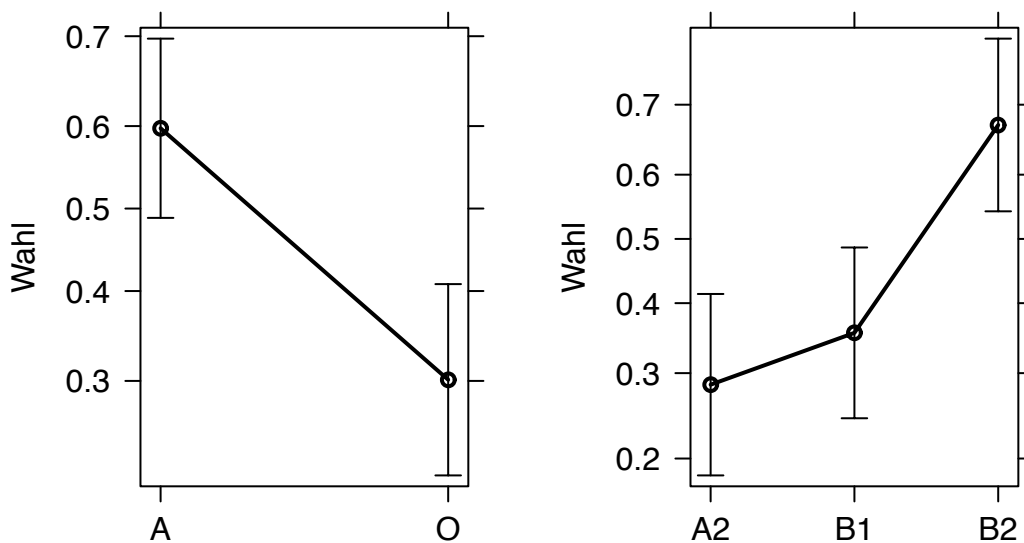


Abbildung 41: Fehltryp Wahl je syntaktischer Funktion der PP und GeR-Gesamtniveaus (Einzeleffekte). Ergebnis des Regressionsmodells RM3a. A = adverbiale PP, O = Objekt-PP.

Die Abbildungen (Abb. 40 und 41) spiegeln somit charakteristische Fehlertypen je syntaktischer Funktion der PP wider, und zwar den Fehlertyp Auslassung für Objekt-PP mit *an/auf* (O) und den Fehlertyp Wahl für adverbiale PP mit *an/auf* (A). Das bedeutet, dass DaF-Lernende die Präpositionen *an/auf* häufiger in Objekt-PP auslassen als in adverbialen PP und dass sie inkorrekte Präpositionen häufiger in adverbialen PP wählen. Die unteren Beispiele illustrieren die beiden Fehlertypen:

Fehlertyp Auslassung:

- *Ich warte Ihrer Antworten* (ID: 1023_0109614)
ZH1: Ich warte auf Ihre Antworten
- *In Freizeit möchte ich viele Aktivitäten teilnehmen* (ID: 1023_0109914)
ZH1: In der Freizeit möchte ich an vielen Aktivitäten teilnehmen

Fehlertyp Wahl:

- *Meine Onkel und Tante leben [...] in die Küste* (ID: 1061_1202917)
ZH1: *Mein Onkel und meine Tante leben [...] an der Küste*
- *ich studiere Informatik auf die Universität* (ID: 1023_0101751)
ZH1: Ich studiere Informatik an der Universität

Die Evaluation der Regressionsmodelle RM2a und RM3a liefert einen Wert von $R^2 = 0.314703$ für den Fehlertyp Auslassung und $R^2 = 0.2157364$ für den Fehlertyp Wahl. Das bedeutet, dass die Regressionsmodelle RM2a und RM3a die Datenverteilung im vorliegenden Datensatz jeweils zu ca. 31% (Fehlertyp Auslassung) bzw. zu 21% (Fehlertyp Wahl) erklären können. Somit zeigt sich ein hohes Erklärungspotenzial der Einzeleffekte der Variablen *Syntaktische Funktion der PP* und *GeR-Gesamtniveau* in ihrem Effekt auf den Fehlertyp (Auslassung und Wahl) im Präpositionsgebrauch durch DaF-Lernende.

Zusammenfassung

Basierend auf den Auswertungen lässt sich die zweite Leitfrage folgendermaßen beantworten:

2. Welche Fehlertypen im Präpositionsgebrauch treten häufig auf und (wie) unterscheiden sich diese gemäß der syntaktischen Funktion der PP (Objekt-PP/adverbiale PP)?

Zu den häufigsten Fehlertypen im Bereich Präpositionsgebrauch (bei *an/auf*-PP) zählen Auslassungen von Präpositionen sowie inkorrekt gewählte Präpositionen. Weiterhin lässt sich feststellen, dass die Wahrscheinlichkeit für Auslassungen mit steigendem GeR-Gesamtniveau sinkt, während die Wahrscheinlichkeit für die inkorrekte Wahl von Präpositionen zunimmt. Dies gilt sowohl für Objekt-PP als auch für adverbiale PP. Gleichzeitig zeigt die statistische Auswertung, dass unabhängig vom GeR-Gesamtniveau der Lernenden ein Einfluss der syntaktischen Funktion auf den Fehlertyp im Präpositionsgebrauch ermittelt werden kann: Auslassungen von Präpositionen sind signifikant wahrscheinlicher in Objekt-PP zu beobachten. Dieses Ergebnis ist auch relevant für die Interpretation der Ergebnisse der ersten Fallstudie zum Einfluss der Erstsprache der Lernenden (s. Kap. 8.5.2.2). Den dortigen Beispielen kann entnommen werden, dass in Kontexten von Objekt-PP realisierungsbedingte Fehler im Präpositionsgebrauch bei Lernenden mit L1TV auftreten. Zur Erinnerung: Unter realisierungsbedingte Fehler fallen inkorrekte hinzugefügte oder ausgelassene Präpositionen. Es ist somit nicht auszuschließen, dass bei realisierungsbedingten Fehlern im Präpositionsgebrauch sowohl der Effekt der L1 als auch der Effekt der syntaktischen Funktion der PP eine Rolle spielt²⁶⁴.

Im Gegensatz zu signifikant höheren Wahrscheinlichkeiten für den Fehlertyp Auslassung bei Objekt-PP können in der vorliegenden Fallstudie in adverbialen PP signifikant höhere Wahrscheinlichkeiten für die inkorrekte Wahl von Präpositionen ermittelt werden. Auch hier lässt sich eine Verbindung zur ersten Fallstudie herstellen. Die analysierten inkorrekten PP-Kontexte der ersten Fallstudie zeigen bei Lernenden mit L1TA vermehrt wahlbedingte Präpositionsfehler in adverbialen PP, Objekt-PP sind dort nur vereinzelt belegt (s. Kap. 8.5.2)²⁶⁵. Auch hier ist anzunehmen, dass sowohl die L1 der Lernenden als auch die syntaktische Funktion der PP einen Einfluss auf den Präpositionsgebrauch bzw. den spezifischen Fehlertyp im Präpositionsgebrauch ausüben. Weiterhin wäre es plausibel

²⁶⁴ Es wäre auch möglich, dass im L1-Subkorpus L1TV insgesamt mehr Kontexte von Objekt-PP enthalten sind als im L1-Subkorpus L1TA und daher u. a. viele Auslassungen von Präpositionen in L1TV festzustellen sind. Die Annahme, dass die Ergebnisse der vorliegenden Fallstudie lediglich auf einen Einfluss ausgewählter L1 der Lernenden zurückzuführen sind, ist dahingegen unplausibel: Die hier analysierten Kontexte von Objekt-PP enthalten insgesamt Daten von Lernenden mit mind. 14 verschiedenen L1. (Der Mindestwert ergibt sich daraus, dass für einige Lernende in MERLIN die L1 nicht dokumentiert ist.)

²⁶⁵ Auch hier ist natürlich nicht auszuschließen, dass insgesamt mehr Kontexte adverbialer PP im Subkorpus L1TA enthalten sein könnten. An dieser Stelle sei ebenfalls erwähnt, dass die hier berichteten Ergebnisse nicht darauf zurückzuführen sind, dass in der Datengrundlage der vorliegenden Fallstudie lediglich einzelne L1 vertreten sind. Die hier analysierten Kontexte von adverbialen PP enthalten insgesamt Daten von Lernenden mit mind. 15 verschiedenen L1.

anzunehmen, dass der Einfluss der syntaktischen Funktion auf den Fehlertyp im Präpositionsgebrauch, der in der vorliegenden Fallstudie exemplarisch für PP-Kontexte mit *an* und *auf* ermittelt wurde, auch bei anderen Präpositionen – z. B. *nach* oder *zu*, die im Datensatz der ersten Fallstudie enthalten sind – greift.

Die statistischen Auswertungen der vorliegenden Fallstudie zum Fehlertyp im Präpositionsgebrauch bestätigen die eingangs formulierten Hypothesen (4./5.) zu spezifischen Fehlertypen je syntaktischer Funktion der PP: Unabhängig vom GeR-Gesamtniveau der Lernenden sind Auslassungen von Präpositionen signifikant wahrscheinlicher in Objekt-PP und wahlbedingte Fehler signifikant wahrscheinlicher in adverbialen PP. Weiterhin zeigt sich im inkorrekten Präpositionsgebrauch unabhängig von der syntaktischen Funktion der PP mit zunehmendem GeR-Gesamtniveau eine Abnahme der Wahrscheinlichkeit für den Fehlertyp Auslassung und eine Zunahme der Wahrscheinlichkeit für den Fehlertyp Wahl, womit die 6. Hypothese bestätigt werden konnte.

9.6 Zusammenfassung und Diskussion der Gesamtergebnisse

In dieser Fallstudie wurde der Einfluss der syntaktischen Funktion der PP auf den Gebrauch von Präpositionen und PP durch DaF-Lernende untersucht. Als Datengrundlage diente das Lernerkorpus MERLIN. In der vorliegenden Fallstudie beschränkte ich mich auf Objekt-PP und adverbiale PP mit den frequenten Präpositionen *auf* und *an*. Ich habe die Annahme überprüft, ob sich die unterschiedlichen Form-Funktionsverbindungen adverbialer PP und Objekt-PP auf den Gebrauch der Präpositionen in entsprechenden PP auswirken bzw. welche Gebrauchsschwierigkeiten sich dahingehend bei DaF-Lernenden auf unterschiedlichen Kompetenzniveaus zeigen. Die Gebrauchsschwierigkeiten wurden durch die Korrektheit und die Fehlertypen im Präpositionsgebrauch operationalisiert. Entsprechend der Fragestellung wurden alle im Korpus vorliegenden PP-Kontexte mit *an/auf* erhoben und manuell aufbereitet, wobei u.a. die syntaktische Funktion der PP und die Fehlertypen im Präpositionsgebrauch (fehler)annotiert wurden. Im Rahmen einer computerunterstützten Fehleranalyse und einer kontrastiven Interlanguage-Analyse wurden die Korrektheit im Präpositionsgebrauch und die Fehlertypen im Präpositionsgebrauch ermittelt und die beiden syntaktischen Funktionen der PP kontrastiert. Die Daten wurden mittels Regressionsanalysen (gemischte Modelle) ausgewertet. Es wurde überprüft, ob bzw. inwiefern die syntaktische Funktion der PP einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch hat. Für die Auswertungen wurde ebenfalls das GeR-Gesamtniveau der Lernenden berücksichtigt. Die Gesamtergebnisse der statistischen Auswertungen zeigen, dass die syntaktische Funktion der PP sowie das GeR-Gesamtniveau der DaF-Lernenden einen Einfluss haben auf den Gebrauch der Präpositionen *an/auf*. Der Einfluss der syntaktischen Funktion auf den Gebrauch der Präpositionen zeigt sich dabei vor allem bei den unterschiedlichen Fehlertypen im Präpositionsgebrauch. Dahingegen scheint die syntaktische Funktion der PP nur einen begrenzten Einfluss auf die Korrektheit im Präpositionsgebrauch auszuüben. In der vorliegenden Fallstudie wurden 379 Objekt-PP und 534 adverbiale PP mit *an/auf* von insgesamt 564 DaF-Lernenden ausgewertet.

Die Ergebnisse der Auswertung zur Korrektheit im Präpositionsgebrauch zeigen zunächst, dass die im Datensatz enthaltenen PP-Kontexte überwiegend korrekt realisiert

werden. Zudem lässt sich beobachten, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch für beide syntaktischen Funktionen (Objekt-PP und adverbiale PP) mit zunehmendem GeR-Gesamtniveau ansteigt. Es erscheint jedoch nicht verwunderlich, dass Lernende vor allem zu Beginn des Fremdspracherwerbs noch nicht über das nötige sprachliche Wissen verfügen, um PP fehlerfrei zu produzieren bzw. Präpositionen (korrekt) zu realisieren. Da dieses Wissen im Verlauf des L2-Erwerbs, d. h. mit Anstieg und Ausbau der sprachlichen Kompetenzen, ebenfalls auf- und ausgebaut wird, nimmt der korrekte Präpositionsgebrauch unabhängig von der syntaktischen Funktion der entsprechenden PP stetig zu.

In der statistischen Auswertung zur Korrektheit im Präpositionsgebrauch kann eine Interaktion zwischen der syntaktischen Funktion der PP und dem GeR-Gesamtniveau der Lernenden ermittelt werden, d. h. der Effekt der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch ist z.T. abhängig vom GeR-Gesamtniveau. Auf dem Niveau der elementaren Sprachverwendung (A2) lässt sich ein starker Effekt der syntaktischen Funktion auf die Korrektheit im Präpositionsgebrauch ermitteln: Auf diesem Kompetenzniveau liegt die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch in adverbialen PP signifikant höher als in Objekt-PP. Auf dem Niveau der selbstständigen Sprachverwendung (B1) lässt sich ein ähnliches Ergebnis ermitteln, der Effekt scheint jedoch deutlich geringer. Auf dem Niveau der höheren Stufe selbstständiger Sprachverwendung (B2) sowie dem Niveau der kompetenten Sprachverwendung (C1) kann jedoch kein Effekt der syntaktischen Funktion der PP auf die Korrektheit im Präpositionsgebrauch ermittelt werden – hier liegen die entsprechenden Wahrscheinlichkeiten für Objekt-PP und adverbiale PP sehr nah beieinander.

Die Annahme, dass Präpositionen innerhalb adverbialer PP aufgrund ihrer prototypischen und zuverlässigen Form-Funktionsverbindungen mit einer höheren Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch einhergehen als Präpositionen in Objekt-PP (1. Hypothese), konnte folglich nur zum Teil bestätigt werden. Jedoch stützt das Ergebnis auf dem Gesamtniveau A2 die Annahme (2. Hypothese), dass zuverlässige Form-Funktionsverbindungen zuerst erworben bzw. beherrscht werden: Adverbiale PP mit *an/auf* zeigen auf diesem Kompetenzniveau nämlich deutlich höhere Wahrscheinlichkeiten für korrekten Präpositionsgebrauch als Objekt-PP mit *an/auf*, die weniger zuverlässige Form-Funktionsverbindungen repräsentieren. Der Erwerb weniger zuverlässiger Form-Funktionsverbindungen erfordert eine höhere kognitive und assoziative Leistung seitens der Lernenden und nimmt folglich eine längere Zeit in Anspruch (Ellis 2002: 153). Die längere Dauer des Erwerbs solcher Form-Funktionsverbindungen könnte sich auch in einer geringeren Wahrscheinlichkeit für korrekten Präpositionsgebrauch in entsprechenden Strukturen widerspiegeln, wie hier im Falle von Objekt-PP mit *an/auf* auf niedrigen Kompetenzniveaus²⁶⁶.

²⁶⁶ Denkbar wäre auch ein Effekt der grammatischen Progression im Sprachunterricht, die Anordnung des zu vermittelnden Lernstoffs in Sequenzen. Möglicherweise werden adverbiale PP im Unterrichtskontext vor Objekt-PP behandelt, sodass Lernende auf einem niedrigen GeR-Gesamtniveau mehr Sicherheit im Gebrauch adverbialer PP zeigen. Eine Durchsicht exemplarischer Übungsgrammatiken für Deutsch als Fremdsprache (Schmidt/Dreyer 2000, Gottschein-Schramm et al. 2017 und Hauschild 2014) zeigt jedoch, dass Präpositionen in Objekt-PP vor Präpositionen in prototypischen adverbialen PP thematisiert werden. Leider liegen mir keine Informationen zu den Lehrwerken vor, mit denen die Lernenden, die in MERLIN repräsentiert sind, gearbeitet haben.

Weiterhin bestätigen die statistischen Auswertungen die 3. Hypothese: Sie zeigen, dass mit zunehmendem Kompetenzniveau sowohl zuverlässige (adverbiale PP) als auch weniger zuverlässige Form-Funktionsverbindungen (Objekt-PP) eine hohe Wahrscheinlichkeit für korrekten Präpositionsgebrauch aufweisen, welches auf einen erfolgreichen Erwerb beider Strukturen hinweist. Mir ist bewusst, dass mittels der Korrektheit nicht direkt auf den Erwerb entsprechender Strukturen geschlossen werden kann (vgl. Baten/Willems 2012: 229f.), jedoch können Auswertungen der Korrektheit Hinweise auf die Beherrschung entsprechender Strukturen liefern und als Tendenzen im Erwerb interpretiert werden (vgl. z. B. Turgay 2011). Damit entsprechen die Ergebnisse der vorliegenden Fallstudie den Tendenzen zum späten Erwerb von Präpositionen in Objekt-PP, die sich ebenfalls bei Bouillon (1977), Diehl et al. (2000) und Harnisch (1993) ableiten lassen (vgl. Kap. 4.1).

Für die Analyse der Fehlertypen im Gebrauch von *an/auf* je syntaktischer Funktion der PP unter Berücksichtigung des GeR-Gesamtniveaus wurden ausschließlich die als inkorrekt annotierten PP-Kontexte ausgewertet. Die Auswertung der Fehlerannotationen zeigt zunächst, dass die Fehlertypen Wahl und Auslassung unabhängig von der syntaktischen Funktion der PP die häufigsten Fehlertypen im Gebrauch von *an/auf* darstellen. Die Ergebnisse der statistischen Auswertungen zeigen, dass die syntaktische Funktion der PP einen signifikanten Effekt auf den Fehlertyp im Präpositionsgebrauch hat und bestätigen damit die 4. und 5. Hypothese: Objekt-PP zeigen eine konstant höhere Wahrscheinlichkeit für Auslassungen von Präpositionen – ein Ergebnis, das sich im Ansatz auch bei Bouillon (1977) und Kuhberg (1990) finden lässt (s. Kap. 4.1). Adverbiale PP dahingegen weisen höhere Wahrscheinlichkeiten für die Wahl einer inkorrekten Präposition auf. DaF-Lernende zeigen somit im Gebrauch der Präpositionen *an* und *auf* in Objekt-PP andere Gebrauchsschwierigkeiten als bei denselben Präpositionen in adverbialen PP: Sowohl die Wahrscheinlichkeit für den Fehlertyp Auslassung als auch den Fehlertyp Wahl unterscheiden sich bei Objekt-PP und adverbialen PP signifikant voneinander, ein Effekt der syntaktischen Funktion der PP auf den Fehlertyp im Präpositionsgebrauch kann hier unabhängig vom GeR-Gesamtniveau ermittelt werden.

Die Ergebnisse zu den Fehlertypen im Präpositionsgebrauch zeigen somit unterschiedliche dominante Fehlertypen, die ihrerseits entsprechende Schwierigkeiten im Gebrauch der Präpositionen *an/auf* reflektieren: Während die Herausforderungen im Bereich Präpositionsgebrauch innerhalb adverbialer PP bei der zielsprachlich korrekten Wahl der jeweiligen Präposition liegen, scheinen die Gebrauchsschwierigkeiten bei Objekt-PP primär mit der Realisierung der jeweiligen Präposition verknüpft zu sein. Die Ergebnisse der statistischen Auswertung liefern somit einen Einblick in die jeweiligen Gebrauchsschwierigkeiten, die DaF-Lernende bei den Präpositionen *an/auf* in spezifischen syntaktischen Strukturen zeigen.

Zugleich bestätigen die statistischen Auswertungen die 6. Hypothese, da sie einen Effekt des GeR-Gesamtniveaus auf den Fehlertyp im Präpositionsgebrauch zeigen. Die Ergebnisse illustrieren, dass die Wahrscheinlichkeiten für den Fehlertyp Auslassung mit steigendem GeR-Gesamtniveau für beide syntaktischen Funktionen der PP sinken, während die Wahrscheinlichkeiten für den Fehlertyp Wahl für beide syntaktischen Funktionen

ansteigen. Das bedeutet, dass Auslassungen von Präpositionen mit Zunahme der Sprachkompetenz insgesamt abnehmen und wahlbedingte Präpositionsfehler mit Zunahme der Sprachkompetenz zunehmen. Da angenommen werden kann, dass wahlbedingte Präpositionsfehler i.d.R. das Wissen um Präpositionen und deren Verwendungskontexte implizieren (was bei Auslassungen i.d.R. nicht angenommen werden kann), spiegelt sich in der Zunahme der Wahrscheinlichkeit für den Fehlertyp Wahl bei gleichzeitiger Abnahme der Wahrscheinlichkeit für den Fehlertyp Auslassung insgesamt auch ein Fortschritt im Erwerb von Präpositionen und PP durch DaF-Lernende wider. Ähnliche Ergebnisse zum Zusammenhang von Fehlertyp und Kompetenzniveau zeigen sich auch bei Meyer-Ingwersen (1975), Keim (1984), Apeltauer (1987), Lütke (2010) und Gutzmann/Turgay (2011) (s. Kap. 4.1).

Zurück zum Einfluss der syntaktischen Funktion auf die Fehlertypen Wahl und Auslassung: Die hohe Wahrscheinlichkeit für Auslassungen in Objekt-PP könnte auf die weniger prototypische, d. h. weniger zuverlässige Form-Funktionsverbindung von Objekt-PP zurückzuführen sein. Eine Objekt-PP erfüllt i.d.R. die Funktion, die parallel ebenfalls durch reine Kasusobjekte angezeigt werden kann²⁶⁷. Für die Realisierung der syntaktischen Funktion ‚Objekt‘ ist die Form einer PP daher nicht prototypisch. Reine Kasusobjekte werden der Kasushierarchie nach prototypisch durch Nominalphrasen (NP) realisiert (vgl. Eisenberg 2013: 63f.). Diese sind frequenter als PP und weisen im Vergleich zu PP einen geringeren Kodierungsaufwand auf (vgl. Kap. 1.2.1). Die hohe Wahrscheinlichkeit präpositionaler Auslassungen in Objekt-PP, die sich in der vorliegenden Fallstudie zeigt, könnte demnach darauf zurückzuführen sein, dass Strukturen ohne Präpositionen (Objekt-NP) frequenter und zuverlässiger für die Kodierung von Objekten sind als Strukturen mit Präpositionen (Objekt-PP). Da zuverlässige und prototypische Form-Funktionsverbindungen früher bzw. schneller erworben werden als weniger zuverlässige und weniger prototypische Form-Funktionsverbindungen, kommt es bei Objekt-PP zur Auslassung der Präposition: Die durch Objekt-PP kodierte syntaktische Funktion wird als einfache NP realisiert. Es ist plausibel anzunehmen, dass diese Eigenschaften den Gebrauch von Präpositionen in Objekt-PP beeinflussen und die Auslassung von Präposition innerhalb solcher PP fördern. Untersuchungen zum Gebrauch von Präpositionen durch Deutschlernende zeigen, dass Auslassungen von Präpositionen charakteristisch sind für niedrige Kompetenzniveaus bzw. wenig entwickelte Lersprachen (s. Kap. 4.1). Dabei wird in diesen Untersuchungen keine Unterscheidung vorgenommen zwischen den verschiedenen syntaktischen Kontexten, in denen Präpositionen ausgelassen werden. Die Daten der vorliegenden Fallstudie bestätigen, dass Auslassungen von Präpositionen eine höhere Wahrscheinlichkeit auf dem Niveau der elementaren Sprachverwendung aufweisen. Allerdings zeigen die Ergebnisse der

²⁶⁷ Breindl (1989: 38f.) merkt jedoch an, dass präpositionale Kasus und reine Kasus auf semantischer Ebene nicht einfach gleichgestellt werden können. Dagegen spräche u.a. die Beobachtung, dass einige Präpositionen innerhalb von Objekt-PP „durchaus bedeutungshaltig“ sein können (vgl. Kap. 1.2.1). Damit wäre folglich eine semantische Gleichsetzung von Objekt-PP und NP nicht immer möglich. Innerhalb von Objekt-PP gehören jedoch die Präpositionen *an/auf* zu den Präpositionen, bei denen „eine Rückführung auf eine Grundbedeutung der Präposition nicht oder nur mit einer etwas gekünstelten Bedeutungshypothese möglich ist“ (Breindl 1989: 39). Zum Einfluss des semantischen Gehalts der Präpositionen in Objekt-PP auf den Gebrauch dieser siehe Kapitel 10.

vorliegenden Fallstudie auch, dass Auslassungen in Objekt-PP auf den mittleren und höheren Kompetenzniveaus zwar abnehmen, jedoch charakteristisch für Objekt-PP bleiben. Dies kann als ein Indiz dafür gesehen werden, dass die Realisierung von Präpositionen in Objekt-PP eine andauernde Herausforderung im DaF-Erwerb und -Gebrauch darstellt.

Neben den weniger zuverlässigen Form-Funktionsverbindungen von Objekt-PP kann auch angenommen werden, dass die Auslassung von *an/auf* in Objekt-PP zusätzlich durch deren semantischen Gehalt bzw. die Bedeutung, die Präpositionen in Objekt-PP einnehmen, begünstigt wird. Diese weicht bei *an/auf* nämlich von der Grundbedeutung dieser Präpositionen ab bzw. ist i.d.R. nicht mehr auf diese zurückführbar (Breindl 1989: 39). Die Beobachtung zu einem möglichen Einfluss des semantischen Gehalts ist ebenfalls kompatibel mit den Annahmen gebrauchsbasierter Ansätze zum L2-Erwerb, die die Bedeutsamkeit von Form-Funktionsverbindungen sowie semantischer Relevanz beim Erwerb sprachlicher Strukturen betonen: Somit stellen Einheiten, die nicht essenziell für die Bedeutung einer Äußerung sind, eine Herausforderung für Lernende dar und werden erst spät erworben (Ellis 2002: 175, Ellis 2006b, Mitchell/Myles/Marsden 2013: 105, s. Kap. 2). Inwiefern der semantische Gehalt einer Präposition in Objekt-PP tatsächlich einen Effekt auf den Gebrauch von Präpositionen durch DaF-Lernende hat, wird in einer separaten Fallstudie in MERLIN systematisch untersucht (Fallstudie 3, Kap. 10).

Die Ergebnisse der vorliegenden Fallstudie liefern empirisch fundierte Erkenntnisse zum Einfluss der syntaktischen Funktion der PP auf den Gebrauch von Präpositionen und PP durch DaF-Lernende. Die Ergebnisse zeigen insgesamt, dass unterschiedliche syntaktische Funktionen von PP (hier: Objekt-PP und adverbiale PP) DaF-Lernende vor unterschiedliche Herausforderungen stellen. Sichtbar werden diese unterschiedlichen Herausforderungen jedoch nicht direkt in der Korrektheit im Präpositionsgebrauch, sondern in den spezifischen Fehlertypen, die sich im Präpositionsgebrauch zeigen. Die Ergebnisse der statistischen Auswertung einer computerunterstützten Fehleranalyse und einer kontrastiven Interlanguage-Analyse zeigen, dass der Fehlertyp Auslassung charakteristisch ist für Präpositionen in Objekt-PP. Dahingegen ist der Fehlertyp Wahl kennzeichnend für Fehler im Gebrauch der Präpositionen *an/auf* innerhalb adverbialer PP. Insgesamt liefern die Ergebnisse der vorliegenden Fallstudie Hinweise darauf, dass in Bezug auf die syntaktische Funktion von PP weniger prototypische bzw. weniger zuverlässige Form-Funktionsverbindungen tendenziell zu mehr, aber vor allem zu anderen Schwierigkeiten im Präpositionsgebrauch führen als prototypische bzw. zuverlässige Form-Funktionsverbindungen.

10. Fallstudie 3: Semantischer Gehalt der Präposition

10.1 Einleitung

In der dritten Fallstudie wird der Einfluss des semantischen Gehalts der Präposition auf den Gebrauch von Präpositionen und PP durch DaF-Lernende untersucht. Hinweise, dass der semantische Gehalt einen Einfluss im Erwerb von Präpositionen und PP ausübt, findet man u.a. bei Carroll/Becker (1993: 146), Hufeisen/Gibson (2002) und Turgay (2010a). Bisher fehlen jedoch systematische, korpusgestützte Untersuchungen in diesem Bereich (vgl. Kap. 4.1). Die

vorliegende Fallstudie leistet einen Beitrag dazu, diese Forschungslücke zu schließen. Im Fokus dieser Fallstudie stehen ausschließlich Objekt-PP (s. Kap. 1.2.1). Untersucht und kontrastiert werden Objekt-PP mit desemantisierten und semantisierten Präpositionen. Als desemantisierte Präpositionen werden die Präpositionen *an/auf* gewählt, als semantisierte Präpositionen die Präpositionen *mit/für*. Wie in den anderen Fallstudien stehen hier die Korrektheit sowie die Fehlertypen im Präpositionsgebrauch im Vordergrund. Aufgrund unterschiedlicher Form-Funktionsverbindungen semantisierter und desemantisierter Präpositionen sowie der damit verbundenen semantischen Relevanz sind nach aktuellen gebrauchsbasierten Ansätzen zum L2-Erwerb Unterschiede im Gebrauch entsprechender Präpositionen zu erwarten. Die vorliegende Fallstudie greift z.T. auf Ergebnisse von Fallstudie 2 (Kap. 9) zurück und verwendet eine Teilmenge der dort erhobenen und aufbereiteten Daten.

10.2 Untersuchungsgegenstand und Leitfragen

Ziel der Fallstudie ist es, zu untersuchen, inwiefern der semantische Gehalt einer Präpositionen in Objekt-PP den Gebrauch dieser Präpositionen durch DaF-Lernende beeinflusst. In dieser Fallstudie wird daher der Gebrauch desemantisierter und semantisierter Präpositionen innerhalb von Objekt-PP gegenübergestellt. Wie aus Kapitel 1.2.1 hervorgeht, lassen sich Präpositionen innerhalb von Objekt-PP nicht allesamt als „bedeutungsneutral“ (Duden 2016: 618) klassifizieren. Vielmehr kann man bei Präpositionen in Objekt-PP nach Breindl (1989: 42) von einem Spektrum der Semantisierung mit den Polen (voll) semantisiert und (voll) desemantisiert ausgehen. Präpositionen in Objekt-PP unterscheiden sich folglich im Grad ihres semantischen Gehalts (Näheres dazu s. Kap. 1.2). Ich beschränke mich hier lediglich auf den semantischen Gehalt von Präpositionen in Objekt-PP, um die syntaktische Funktion der PP konstant zu halten. Ich unterscheide zwischen *desemantisierten* und *semantisierten* Präpositionen. Als desemantisiert gelten hier die Präpositionen *an/auf*, als semantisiert die Präpositionen *mit/für*. Damit ist jedoch nicht ausgesagt, dass die semantisierten Präpositionen *mit/für* immer eine eindeutig isolierbare Bedeutung in Objekt-PP aufweisen und für die desemantisierten Präposition *an/auf* niemals ein Zusammenhang zur Grundbedeutung hergestellt werden kann²⁶⁸. Ich gehe hier von prototypischen, d. h. häufigen Vorkommnissen aus und orientiere mich hierfür an den Ausführungen in Breindl (1989: 39ff.). Zur Erinnerung sei an dieser Stelle (Tab. 57) der unterschiedliche semantische Gehalt sowie die unterschiedlichen Form-Funktionsverbindungen semantisierter und desemantisierter Präpositionen erneut aufgeführt (s. auch Kap. 1.2.1).

²⁶⁸ Zu den Verwendungen der Präpositionen *an/auf* sowie *mit/für* und ihre Nähe zur Grundbedeutung der entsprechenden Präpositionen siehe z. B. GDS (1997: 2116ff.).

Form: Präposition	Funktion: Semantischer Gehalt (Objekt-PP)	Form-Funktionsverbindung
<i>mit</i> <i>für</i>	semantisiert	prototypisch und zuverlässig
<i>an</i> <i>auf</i>	desemantisiert	nicht prototypisch und weniger zuverlässig

Tabelle 57: Semantisierte und desemantisierte Präpositionen (in Objekt-PP): Form-Funktionsverbindungen.

In Bezug auf den Erwerb deutscher Präpositionen und PP ist aus der Perspektive aktueller gebrauchsbasierter Ansätze zum L2-Erwerb davon auszugehen, dass der Unterschied in der Zuverlässigkeit der Form-Funktionsverbindungen desemantisierter und semantisierter Präpositionen einen Einfluss auf den Erwerb der entsprechenden Präpositionen in Objekt-PP hat. Es wird davon ausgegangen, dass prototypische und zuverlässige Form-Funktionsverbindungen im Spracherwerb leichter und früher erworben werden als weniger zuverlässige Form-Funktionsverbindungen (vgl. MacWhinney 2001, Ellis 2003) (s. Kap. 2). Diese grundlegende Annahme hat ebenfalls die Fallstudie zum Einfluss der syntaktischen Funktion der PP auf den Gebrauch von Präpositionen und PP geleitet (Fallstudie 2, Kap. 9). Dort wurden die Form-Funktionsverbindungen syntaktischer Strukturen – PP – in den Fokus gerückt. Form-Funktionsverbindungen sind jedoch auf unterschiedlichen sprachlichen Ebenen evident, u.a. auch auf der Ebene einzelner Lexeme (Ellis/Wulff 2015: 76). Betrachtet man nun die Form-Funktionsverbindungen von Präpositionen in Objekt-PP (s. Tab. 57), sind Unterschiede im Gebrauch desemantisierter und semantisierter Präpositionen zu erwarten. Weiterhin greift der Einfluss semantischer Relevanz sprachlicher Formen auf den Erwerb dieser. Wie Wegener (1995c: 3) deutlich macht, ist

[a]us psycholinguistischen Untersuchungen [...] bekannt, daß Lerne[nde] Formen, die semantische und/oder grammatische Funktionen anzeigen, leichter lernen als solche, die arbiträr sind oder die [den Lernenden] zumindest vorläufig als funktionslos erscheinen.

Da desemantisierte Präpositionen innerhalb von Objekt-PP in ihrer Bedeutung nicht mehr (gut) isolierbar sind, ist anzunehmen, dass sie folglich keine semantische Funktion in Objekt-PP aufweisen und damit mehr Gebrauchsschwierigkeiten erwarten lassen als semantisierte Präpositionen, die in ihrer Bedeutung gut isolierbar sind und für die somit eine semantische Funktion, und damit verbunden eine semantische Relevanz, abgeleitet werden kann²⁶⁹. Die Fallstudie untersucht den Einfluss des semantischen Gehalts der Präposition auf den Präpositionsgebrauch. Die Leitfragen der Fallstudie lauten wie folgt:

1. Welche Unterschiede zeigen sich bei der Korrektheit im Gebrauch semantisierter Präpositionen im Vergleich zu desemantisierten Präpositionen in Objekt-PP?

²⁶⁹ Hufeisen/Gibson (2002) können entsprechende Tendenzen im Präpositionsgebrauch bei DaF-Lernenden feststellen.

2. Welche Fehlertypen treten bei *semantisierten Präpositionen* häufig auf und (wie) unterscheiden sich diese von den Fehlertypen im Gebrauch von *desemantisierten Präpositionen* in Objekt-PP?

Bei beiden Leitfragen wird jeweils zusätzlich das GeR-Gesamtniveau der Lernenden berücksichtigt.

Entsprechend der obigen Ausführungen ist für die vorliegende Fallstudie anzunehmen, dass DaF-Lernende im Gebrauch semantisierter Präpositionen (*mit/für*) in Objekt-PP weniger Schwierigkeiten aufzeigen als im Gebrauch desemantisierter Präpositionen (*an/auf*). Entsprechend den anderen Fallstudien werden Gebrauchsschwierigkeiten hier durch die Korrektheit und Fehlertypen im Präpositionsgebrauch operationalisiert. Den Annahmen gebrauchsbasierter Ansätze zum L2-Erwerb folgend gehe ich in Bezug auf den Gebrauch semantisierter und desemantisierter Präpositionen in Objekt-PP von folgenden Hypothesen aus, die im Rahmen der Fallstudie korpusgestützt überprüft werden:

Hypothesen zu Leitfrage 1:

1. Die Korrektheit im Präpositionsgebrauch ist insgesamt höher bei semantisierten Präpositionen im Vergleich zu desemantisierten Präpositionen.
2. Unabhängig vom semantischen Gehalt der Präposition nimmt die Korrektheit im Präpositionsgebrauch mit Anstieg des Kompetenzniveaus zu.

Hypothesen zu Leitfrage 2:

3. Desemantisierte Präpositionen werden häufiger ausgelassen als semantisierte Präpositionen.
4. Desemantisierte und semantisierte Präpositionen werden ähnlich häufig inkorrekt gewählt bzw. in Kontexten hinzugefügt, in denen zielsprachlich keine Präpositionen gefordert sind.
5. Unabhängig vom semantischen Gehalt der Präposition nehmen Auslassungen von Präpositionen mit Anstieg des Kompetenzniveaus ab, während wahlbedingte Präpositionsfehler zunehmen.

10.3 Datenauswahl und Datenerhebung

Äquivalent zu Fallstudie 2 sind für diese Fallstudie im Lernerkorpus MERLIN vorrangig die Metadaten der Lernertexte (Identifikationsnummer, GeR-Gesamtniveau), die Ebene der Zielhypothese (ZH1) sowie die Fehlerannotationskategorien im Bereich Präpositionsgebrauch relevant. Die Datenauswahl orientiert sich am Untersuchungsgegenstand, d. h. der Gegenüberstellung des Gebrauchs desemantisierter und semantisierter Präpositionen in Objekt-PP.

Ich beschränke mich auf Objekt-PP, da die Unterscheidung zwischen semantisierten und desemantisierten Präpositionen besonders in dieser syntaktischen Funktion interessant ist. Wenn Breindl (1989: 39) schreibt, dass bei einigen Präpositionen innerhalb von Objekt-PP „eine Rückführung auf eine Grundbedeutung der Präpositionen“ nicht gut möglich ist, dann impliziert dies, dass die Grundbedeutung solcher Präpositionen außerhalb von Objekt-PP zu

suchen ist. Abhängig vom semantischen Gehalt der Präposition ist somit die Abweichung zwischen der Präpositionsbedeutung innerhalb und außerhalb von Objekt-PP unterschiedlich. Das Wirken dieser ‚semantischen Abweichung‘ lässt sich somit primär in Objekt-PP untersuchen. Des Weiteren unterstützen die Ergebnisse aus Fallstudie 2 die Fokussierung auf eine einzige syntaktische Funktion: Jene Ergebnisse zeigten, dass die syntaktische Funktion einer PP einen Einfluss auf den Gebrauch (Korrektheit und Fehlertyp) von Präpositionen und PP hat. Durch die Beschränkung auf eine einzige syntaktische Funktion (hier: Objekt-PP) kann die Variable *Semantischer Gehalt der Präposition* gezielter untersucht und ein zusätzlicher Einfluss der syntaktischen Funktion ausgeschlossen werden.

Den Ausführungen von Breindl (1989) folgend werden als semantisierte Präpositionen in Objekt-PP die Präpositionen *mit* und *für* ausgewählt (s. auch Kap. 1.2.1). Zur Erinnerung sei an dieser Stelle erwähnt, dass diese Präpositionen „relativ konstante Bedeutungszüge“ (Breindl 1989: 39) unabhängig von der syntaktischen Funktion der PP zeigen. Damit weisen sie recht zuverlässige Form-Funktionsverbindungen auf. Im Gegensatz dazu werden als desemantisierte Präpositionen die Präpositionen *an* und *auf* ausgewählt. In Objekt-PP ist bei diesen Präpositionen nach Breindl (1989: 39) „eine Rückführung auf eine Grundbedeutung der Präpositionen nicht oder nur mit einer etwas gekünstelten Bedeutungshypothese möglich“. Diese Präpositionen weisen damit keine zuverlässigen bzw. weniger zuverlässige Form-Funktionsverbindungen auf. Für die vorliegende Fallstudie werden PP-Kontexte²⁷⁰ mit den Präpositionen *mit* und *für* im MERLIN-Korpus erhoben, aufbereitet und ausgewertet. Ich gehe weiter unten näher darauf ein. PP-Kontexte mit den Präpositionen *an* und *auf* liegen mir bereits aus Fallstudie 2 vor. Diese werden für die Datenauswertung der vorliegenden Fallstudie herangezogen, um den lernersprachlichen Gebrauch semantisierter (*mit/für*) und desemantisierter (*an/auf*) Präpositionen in Objekt-PP gegenüberzustellen. Die vorliegende Fallstudie greift somit auf eine Teilmenge der erhobenen und aufbereiteten Korpusdaten aus Fallstudie 2 zurück.

Für die Datenerhebung der Präposition *mit* und *für* in MERLIN gehe ich parallel zu Fallstudie 2 vor: Die Abfrage im Lernerkorpus erfolgt wortformbasiert, wobei die Groß- und Kleinschreibung gleichermaßen berücksichtigt wird. Verschmelzungen aus Präposition und Artikel werden bei der Korpusabfrage ebenfalls berücksichtigt. Die Abfrage erfolgt, parallel zum Vorgehen von Fallstudie 2, auf Ebene der ZH1. Mit einer Abfrage auf dieser Ebene können alle PP-Kontexte ermittelt werden, in denen die Verwendung einer *mit*- bzw. *für*-PP zielsprachlich gefordert ist. Somit werden mit der wortformbasierten Korpusabfrage auf Ebene der ZH1 alle PP-Kontexte ermittelt, in denen die entsprechende Präposition korrekt, inkorrekt oder gar nicht realisiert wurde. Um ebenfalls PP-Kontexte zu erheben, in denen DaF-Lernende die Präpositionen *mit* und *für* realisiert haben, diese jedoch zielsprachlich nicht gefordert waren (und daher auf Ebene der ZH1 getilgt wurden), beziehe ich mich in einer zusätzlichen Abfrage auf die ZH1Diff-Ebene und das Editiertag DEL (Näheres s. Kap. 3.3.2).

²⁷⁰ Mit dem Begriff *PP-Kontexte* meine ich sowohl den zielsprachlichen als auch den nicht-zielsprachlichen Gebrauch von Präpositionen bzw. PP. Der nicht-zielsprachliche Gebrauch inkludiert dabei auch Kontexte, in denen der Lernende keine Präposition bzw. PP realisiert hat, zielsprachlich jedoch eine Präposition bzw. PP gefordert ist (vgl. Kap. 9.3).

Durch die Kombination aus spezifischer Wortform (in diesem Fall *mit* und *für*²⁷¹) und dem Editiertag DEL lassen sich die PP-Kontexte inkorrekt Hinzufügungen der ausgewählten Präpositionen im Korpus erheben. Durch den Bezug auf die Ebene ZH1- bzw. ZH1-Diff werden somit alle im Korpus enthaltenen PP-Kontexte der ausgewählten Präpositionen *mit* und *für* erfasst. Wie auch in Fallstudie 2 verzichte ich hier auf die reine Abfrage fehlerannotierter PP-Kontexte, da in MERLIN nicht alle Lernertexte fehlerannotiert sind (s. Kap. 9.3).

Die Korpusabfragen der vorliegenden Fallstudie sind in der folgenden Tabelle (Tab. 58) abgebildet, die Ergebnisse beider Abfragen sind in Tabelle 59 je Präposition und GeR-Gesamtniveau zusammengefasst²⁷². Da die GeR-Gesamtniveaus A1 und C2 in Fallstudie 2 nicht für die Datenauswertung berücksichtigt wurden, werden die entsprechenden GeR-Gesamtniveaus für die hier vorliegende Fallstudie ebenfalls nicht beachtet. Aus diesem Grund sind die Niveaus A1 und C2 in Tabelle 59 nicht aufgeführt.

Korpusabfrage	Beschreibung
TH1=/[Ff]ür [Ff]ürs [Ff]ürn ²⁷³ / TH1=/[Mm]it/ Abfrage je GeR-Gesamtniveau	Abfrage aller Instanzen entsprechender Wortformen in zielsprachlich geforderten Kontexten = Ermittlung von PP-Kontexten, in denen die Präposition auf Ebene der ZH1 nicht getilgt wurde.
tok=/[Ff]ür [Ff]ürs [Ff]ürn/_=_TH1Diff="DEL" tok=/[Mm]it/_=_TH1Diff="DEL" Abfrage je GeR-Gesamtniveau	Abfrage aller Instanzen entsprechender Wortformen in zielsprachlich nicht geforderten Kontexten = Ermittlung von PP-Kontexten, in denen die Präposition auf Ebene der ZH1 getilgt wurde.

Tabelle 58: Korpusabfragen zur Fallstudie Semantischer Gehalt der Präposition.

Präposition (Wortform)	Anzahl Treffer im Korpus							
	GeR-Gesamtniveau							Gesamt
	A2	A2+	B1	B1+	B2	B2+	C1	
<i>mit</i>	181	103	252	218	465	141	76	1436
<i>für</i>	118	89	269	176	367	131	108	1258
Summe	299	192	521	394	832	272	184	2694

Tabelle 59: Ergebnis der Korpusabfragen zur Fallstudie Semantischer Gehalt der Präposition (Wortform) je Präposition und GeR-Gesamtniveau (nicht bereinigt).

²⁷¹ Fortan sind bei Nennung der Präpositionen *mit* und *für* ebenfalls deren Verschmelzungen inbegriffen.

²⁷² Die Tabelle enthält die Trefferanzahl der jeweils abgefragten Wortform *mit/für*. Die Werte sind nicht bereinigt, die Tabelle enthält u.a. noch Pseudotreffer, die in der Phase der Datenaufbereitung aussortiert werden.

²⁷³ Für die Verschmelzungsform *fürn/Fürn* konnte kein Korpusstreffer ermittelt werden. Sie ist hier der Vollständigkeit halber dennoch angegeben.

Insgesamt liefern die Korpusabfragen in MERLIN 2694 Treffer. Die Treffer werden mit dem im Korpusrecherchesystem ANNIS integrierten GridExporter getrennt nach Präposition und GeR-Gesamtniveau exportiert. Je Treffer werden für einen Kontext von zehn Token rechts und links folgende Daten exportiert: Die Original-Lerneräußerung, die ZH1-Ebene und ggf. ZH1Diff-Ebene. Ebenfalls exportiert werden die Fehlerannotationen im Bereich Präpositionsgebrauch, soweit vorliegend. Diese werden bei der Auswertung zur Korrektheit und den Fehlertypen im Präpositionsgebrauch berücksichtigt. Als Metadaten werden die Erstsprache (L1) des/der Lernenden und die Identifikationsnummer (kurz ID) exportiert. Die exportierten Korpusstreffer werden getrennt nach GeR-Gesamtniveau und Präposition in Excel-Dokumente importiert und dort für die Datenauswertung aufbereitet.

10.4 Datenaufbereitung

In diesem Kapitel wird ausschließlich die Datenaufbereitung der PP-Kontexte mit den Präpositionen *mit/für* näher beschrieben. Zuerst gehe ich auf die Annotationskategorien ein und beschreibe im Anschluss den Annotationsprozess. Das Vorgehen entspricht dem von Fallstudie 2 (s. Kap. 9.4).

10.4.1 Annotationskategorien

Die exportierten Korpusstreffer werden getrennt nach Präposition und GeR-Gesamtniveau in einzelnen Excel-Dokumenten aufbereitet. Da ich in der vorliegenden Fallstudie auf eine Teilmenge aufbereiteter Daten aus Fallstudie 2 zurückgreife, muss sichergestellt werden, dass die Datenaufbereitung der Korpusstreffer in der vorliegenden Fallstudie einem identischen Vorgehen folgt. Aus diesem Grund setze ich in der vorliegenden Fallstudie dieselben Annotationskategorien an wie in der zweiten Fallstudie (s. Kap. 9.4.). Im Rahmen der Datenaufbereitung sollen primär Pseudotreffer aussortiert und die PP-Kontexte hinsichtlich der syntaktischen Funktion der PP kategorisiert werden. Das Annotationshandbuch, das im Rahmen von Fallstudie 2 erstellt wurde, wird ebenfalls für die Datenaufbereitung in der vorliegenden Fallstudie eingesetzt. Das Handbuch wurde hierfür um Beispiele und Hinweise zur Annotation von PP-Kontexten mit den Präpositionen *mit* und *für* ergänzt. Die Annotationskategorien und -richtlinien sind im Annotationshandbuch ausführlich beschrieben (s. Anhang). Zur besseren Nachvollziehbarkeit seien die Annotationskategorien an dieser Stelle wiedergegeben (Tab. 60) und hier an Beispielen mit den Präpositionen *mit* und *für* illustriert²⁷⁴.

²⁷⁴ In der vorliegenden Fallstudie stehen lediglich Objekt-PP im Vordergrund. Der Vollständigkeit halber werden jedoch im Rahmen der Datenaufbereitung in der vorliegenden Fallstudie auch andere syntaktische Funktionen von PP annotiert. Diese Annotationen stehen somit für eine weitere Auswertung bei Bedarf zur Verfügung.

Annotationskategorie (Annotationswert 1/0) ²⁷⁵		Erläuterung	Beispiele (z.T. aus MERLIN ²⁷⁶)
Syntaktische Funktion der PP	1	Objekt-PP	PP in Objekt-Funktion. Hierzu zähle ich auch prädikative PP und Adjektive mit Präpositionsreaktion
	2	Adverbiale PP	PP in adverbialer Funktion
	3	Attributive PP	PP in attributiver Funktion
Weitere Annotationskategorien	4	Rest	PP-Kontexte, deren Zuordnung in andere Kategorien aufgrund unterschiedlicher Gründe nicht möglich war.
	5	Pseudotreffer	Form ist keine PP, sondern z. B. eine Verbpartikel oder eine substantivierte Präposition

²⁷⁵ Die Annotationswerte 1 und 0 entsprechen der Zuordnung (1) bzw. Nicht-Zuordnung (0) einer PP in die entsprechende Annotationskategorie. Näheres s. Annotationshandbuch im Anhang.

²⁷⁶ Die Beispiele repräsentieren Original-Lerneräußerungen. Zur besseren Nachvollziehbarkeit ist die ZH1 vereinzelt angegeben.

²⁷⁷ In diesem Beispiel wurde die *für*-PP in der Original-Lerneräußerung auf Ebene der ZH1 durch eine Nominalphrase im Nominativ ersetzt. Die Nominalphrase im Nominativ hat hier Subjektfunktion. Da die syntaktische Funktion *Subjekt* in der vorliegenden Fallstudie nicht behandelt wird, wird solch ein PP-Kontext als *Rest* annotiert. Nähere Informationen zur Annotation von inkorrekt hinzugefügten Präpositionen finden sich im Annotationshandbuch (s. Anhang).

²⁷⁸ In diesem Beispiel wurde die *mit*-PP (*mit mein Arbeit*) durch eine Nominalphrase ersetzt (*meiner Arbeit*) und *nähe* zu *nahe* geändert. Auf Ebene der ZH1 liegt die PP *nahe meiner Arbeit* vor. Somit wäre die *mit*-PP in der Original-Lerneräußerung als Bestandteil einer anderen PP (*nahe mit mein Arbeit*) zu interpretieren.

	6	Nicht interpretierbar	PP-Kontext (auf Ebene der ZH1) aufgrund von z. B. semantischen Abweichungen nicht interpretierbar bzw. unverständlich ²⁷⁹	<i>ich könnte nicht an meine Kenntnisse mit euren teilnehmen</i> ZH1: [...] <i>ich könnte nicht an meine Kenntnisse mit euren teilnehmen</i> ²⁸⁰ . <i>Bitte, schön für mein Herz, du fahre.</i> ZH1: <i>Bitte, schön für mein Herz, du fährst.</i> ²⁸¹
	7	Funktionsverbgefüge ²⁸²	Prädikatäquivalent aus Funktionsverb und PP	<i>Zur Anklage kommen</i> <i>In Anspruch nehmen</i>
	8	Formelhafte Sequenz	PP, die formelhafte Sequenzen darstellen (s. auch Kap. 1.2.5)	<i>Mit freundlichen Grüßen</i> <i>Vielen Dank für [...]</i>

Tabelle 60: Annotationskategorien in der Fallstudie Semantischer Gehalt der Präposition.

Bei Objekt-PP werden im Rahmen der Annotation zusätzlich die entsprechenden Verben bzw. Adjektive erfasst, das Vorgehen ist im Annotationshandbuch beschrieben (s. Anhang). Zur Beantwortung der Leitfragen dieser Fallstudie sind lediglich Kontexte von Objekt-PP von Relevanz. An dieser Stelle sei daher bereits erwähnt, dass diejenigen PP-Kontexte mit *mit/für*, die anderen Annotationskategorien zugeordnet werden, von der weiteren Auswertung ausgeschlossen sind. Zur Bestimmung von Objekt-PP (und zur Abgrenzung dieser von adverbialen PP) werden, wie auch in Fallstudie 2, u.a. linguistische Ressourcen (z. B. E-Valbu oder Duden-Online) konsultiert (Näheres s. Annotationshandbuch im Anhang).

Bei der Erfassung von Fehlern und der Annotation der Fehlertypen im Präpositionsgebrauch gehe ich ebenfalls parallel zu Fallstudie 2 vor. Das Verfahren möchte ich an dieser Stelle nochmals kurz skizzieren: Wie auch in den anderen Fallstudien definiere ich einen Fehler im Präpositionsgebrauch als eine Abweichung zwischen der Original-Lerneräußerung und der entsprechenden Zielhypothese (ZH1), wenn diese Abweichung die Präposition selbst betrifft. Zu den Fehlern im Präpositionsgebrauch zählen primär Ersetzungen, Löschungen, Ergänzungen oder Positionsänderungen der Präpositionen *mit/für*

²⁷⁹ Da auf Ebene der ZH1 lediglich grammatische und orthographische Fehler korrigiert sind, ist es möglich, dass die ZH1 semantische Abweichungen enthält, die erst auf Ebene der ZH2 korrigiert werden (vgl. MERLIN project 2014). In der vorliegenden Fallstudie beziehe ich mich lediglich auf die Ebene ZH1.

²⁸⁰ Bei diesem Beispiel handelt es sich um einen ungrammatischen Satz auf Ebene der ZH1.

²⁸¹ Bei diesem Beispiel ist die Aussage unklar, zumal auf ZH1-Ebene die Äußerung nur minimal geändert wurde.

²⁸² Ich möchte hier anmerken, dass im Datensatz der vorliegenden Fallstudie keine PP-Kontexte von FVG vorliegen.

auf Ebene der ZH1. Derartige Änderungen repräsentieren *inkorrekte* PP-Kontexte. Unter *korrekten* PP-Kontexten verstehe ich im Umkehrschluss solche PP-Kontexte, in denen keine Änderung zwischen Original-Lerneräußerung und ZH1 in Bezug auf die Präposition vorliegt. Diese Fehlerdefinition bildet in der vorliegenden Fallstudie die Grundlage für die Annotationen und die Auswertungen der Fehlertypen im Bereich Präpositionsgebrauch (s. Kap. 10.5).

Für die Erfassung inkorrektur PP-Kontexte sowie die Annotation der Fehlertypen im Präpositionsgebrauch orientiere ich mich an den Fehlerannotationskategorien im Bereich Präpositionsgebrauch in MERLIN. Aus diesem Grund werden, entsprechend Fallstudie 2, als Annotationswerte für die einzelnen Fehlertypen die Kürzel eingefügt, die sich ebenfalls im Annotationsschema von MERLIN finden (vgl. dazu auch Wisniewski et al. 2014: 12). Zur besseren Übersicht sind die Annotationskategorien zum Fehlertyp im Präpositionsgebrauch sowie die einzelnen Annotationswerte in der nachfolgenden Tabelle (Tab. 61) aufgeführt. Beispiele für entsprechende PP-Kontexte mit den Präpositionen *mit/für* sind jeweils angegeben²⁸³.

Annotationskategorie <i>Fehlertyp</i> (Annotationswert)		Beispiel (aus MERLIN)
1	Kein Fehler im Präpositionsgebrauch (0)	<i>[...] in zwei Monaten muss ich mit der neuen Job anfangen.</i> ZH1: <i>[...] in zwei Monaten muss ich mit dem neuen Job anfangen.</i>
2	Auslassung einer Präposition (o)	<i>Ich interessiere mich sehr das Wohnung.</i> ZH1: <i>Ich interessiere mich sehr für die Wohnung</i>
3	Hinzufügung einer Präposition (ad)	<i>Ich mag auch natürlich mit anderen Gebräuchen kennen lernen.</i> ZH1: <i>Ich mag auch natürlich andere Gebräuche kennenlernen.</i>
4	Wahl der Präposition (ch)	<i>Ich interessiere mich über elektrotechnik.</i> ZH1: <i>Ich interessiere mich für Elektrotechnik.</i>
5	Position der Präposition (pos)	<i>[...] Informativ Schwerpunkt mit Systemkaufmann.</i> ZH1: <i>[...] Informatik mit Schwerpunkt Systemkaufmann.</i>
6	Zusammenziehung der Präposition (merge)	<i>wie viele Stunden Prowoche [...]</i> ZH1: <i>Wie viele Stunden pro Woche [...]</i>
7	Trennung der Präposition (split)	<i>Da ich zur Zeit keine Familie habe [...]</i> ZH1: <i>Da ich zurzeit keine Familie habe [...]</i>

Tabelle 61: Annotationskategorien und -werte zum Fehlertyp im Bereich Präpositionsgebrauch in der Fallstudie Semantischer Gehalt der Präposition.

²⁸³ Die Beispiele unter 6 und 7 (inkorrekte Zusammenziehung bzw. inkorrekte Trennung einer Präposition) sind identisch mit den Beispielen aus der entsprechenden Tabelle aus Fallstudie 2. Dies liegt daran, dass diese Fehlertypen im Datensatz der vorliegenden Fallstudie nicht vertreten sind.

Für die vorliegende Fallstudie sind vor allem die Fehlertypen 1.-5. relevant. Diese stehen ebenfalls in Fallstudie 2 im Fokus (s. Kap. 9.4). Die vorhandenen exportierten Fehlerannotationen im Bereich Präpositionsgebrauch im Lernerkorpus MERLIN werden gesichtet und, wo notwendig, im Rahmen der Datenaufbereitung manuell korrigiert. Nicht vorhandene Fehlerannotationen werden nachträglich manuell ergänzt. Grundlage der nachträglich ergänzten Fehlerannotationen bildeten das Annotationsschema und die Annotationsrichtlinien des MERLIN-Lernerkorpus. Die Annotationskategorien und -richtlinien zur Annotation der Fehlertypen im Präpositionsgebrauch sind ebenfalls im Annotationshandbuch beschrieben (s. Anhang). Für die Datenauswertung werden die korrigierten und die nachträglich ergänzten Fehlerannotationen berücksichtigt. Da die ZH1 für alle Lernertexte im Korpus vorliegt, ist eine konsistente Fehleridentifikation und -annotation gewährleistet.

10.4.2 Annotationsprozess

Die exportierten PP-Kontexte mit den Präpositionen *für* und *mit* werden nach den erwähnten Annotationskategorien (s. Tab. 60) durch zwei studentische Hilfskräfte (SHK1 und SHK2²⁸⁴) mit Rückgriff auf das Annotationshandbuch (s. Anhang) manuell annotiert. SHK1 wurde bereits im Rahmen von Fallstudie 2 für Annotationen eingesetzt. Eine andere Hilfskraft (SHK2) wurde von ihr in den Annotationsprozess und das Annotationshandbuch eingearbeitet. Zu Beginn der Phase der Datenaufbereitung in der vorliegenden Fallstudie wurde die Annotation der PP-Kontexte mit *mit/für* auf Basis des erweiterten Annotationshandbuchs im Rahmen von Probeannotationen durch beide SHK eingeübt.

Zur Ermittlung der Reliabilität der Annotationen (gemäß den Annotationskategorien aus Tab. 60) habe ich das Inter-Annotator-Agreement (kurz IAA) zwischen SHK1, SHK2 und mir berechnet. Für die Berechnung des IAA wurden randomisiert computergestützt je 50 PP-Kontexte mit den Präpositionen *mit* und *für* (also insgesamt 100 PP-Kontexte) aus den GeR-Gesamtniveaus A2-C1 ausgewählt. Die einzelnen GeR-Gesamtniveaus waren dabei zu gleichen Anteilen repräsentiert. Es wurde sichergestellt, dass die 100 PP-Kontexte keine Pseudotreffer enthielten. Zur Berechnung des IAA wurde entsprechend dem Vorgehen von Fallstudie 2, das Kappa-Maß (κ) gewählt (vgl. Cohen 1960, Artstein/Poesio 2008: 559ff., Krippendorff 1980: 135) (Näheres s. Kap. 9.4.2). Zur Erinnerung: Ein κ -Wert von 0 entspricht keiner Übereinstimmung und ein κ -Wert von 1 zeigt eine vollständige Übereinstimmung zwischen AnnotatorInnen an.

Die Berechnung des κ -Werts erfolgte in der vorliegenden Fallstudie in mehreren Schritten: 1) IAA zwischen SHK1 und mir, 2) IAA zwischen SHK2 und mir und 3) IAA zwischen SHK1 und SHK2. Die Aufteilung der IAA-Berechnungen war notwendig, da SHK1 und SHK2 nicht parallel, sondern nacheinander Annotationen angefertigt haben.

²⁸⁴ Die Aufteilung der Annotationen auf zwei SHK war erforderlich, da SHK1 ihr Studium beendete und die Annotationen zu diesem Zeitpunkt noch nicht abgeschlossen waren.

Für die Berechnung des IAA zwischen SHK1 und mir wurden die Annotationen aller Annotationskategorien²⁸⁵ berücksichtigt (s. Tab. 60). Das IAA lieferte einen κ -Wert von ca. 0.53²⁸⁶ und damit kein zufriedenstellendes Ergebnis. Im Anschluss daran wurde das Annotationshandbuch überarbeitet und u.a. weitere Hinweise zur Annotation von PP-Kontexten mit *mit/für* ergänzt. Im Anschluss wurden die Annotationen überprüft und ggf. korrigiert und das IAA erneut berechnet. Die zweite Runde der IAA-Berechnung lieferte einen ähnlichen Wert wie die erste ($\kappa = 0.52$). Eine genaue Durchsicht der Annotationen, die SHK1 angefertigt hatte, zeigte jedoch, dass eine Mehrzahl der Annotationsabweichungen zwischen SHK1 und mir auf formale Fehler in den Annotationen von SHK1 zurückgeführt werden konnte. So wurden Hinweise zur Vergabe von Annotationswerten im Annotationshandbuch nicht befolgt bzw. übersehen und als Folge nicht während der IAA-Annotationen angewandt. Nach einer Korrektur der formalen Fehler konnte ein κ -Wert von $\kappa = 0.70$ ermittelt werden. Dieser Wert ist insgesamt als zufriedenstellend anzusehen. Im Anschluss an die Berechnung des IAA wurden in einem anschließenden Gespräch zwischen SHK1 und mir alle Abweichungen und unsichere Fälle intensiv besprochen und das Annotationshandbuch entsprechend überarbeitet.

Die Berechnung des IAA zwischen SHK2 und mir erfolgte zu einem späteren Zeitpunkt, als SHK2 als Annotatorin für die vorliegende Fallstudie eingesetzt wurde. Die doppelte Annotation der 100 PP-Kontexte mit *mit/für* erfolgte nach einer Phase der Probeannotation, in der sich SHK2 mit der Annotationsaufgabe, den vorliegenden Daten sowie dem damals aktuellen Annotationshandbuch vertraut machte. Für die Berechnung des IAA wurden dieselben 100 PP-Kontexte gewählt, die bereits für die Berechnung des IAA zwischen SHK1 und mir annotiert wurden. SHK2 annotierte die 100 PP-Kontexte zum ersten Mal. Ich sichtete meine Annotationen der entsprechenden 100 PP-Kontexte erneut und passte sie an den aktuellen Stand des Annotationshandbuchs an, damit ein systematischer Vergleich der Annotationen von SHK2 und mir möglich war. Für das entsprechende IAA wurde ein κ -Wert von $\kappa = 0.72$ ermittelt. Es handelt sich um ein zufriedenstellendes Ergebnis.

Um die Übereinstimmung der (zeitlich versetzten) Annotationen von SHK1 und SHK2 zu ermitteln, habe ich die Annotationen der 100 ausgewählten PP-Kontexte mit *mit/für* von SHK1 denen von SHK2 gegenübergestellt. Hier war jedoch ein vollständiger Vergleich aller Annotationskategorien (s. Tab. 60) nicht möglich. Dies hängt u.a. damit zusammen, dass das Annotationshandbuch zwischenzeitlich überarbeitet wurde und die Annotationen der 100 *mit/für*-PP-Kontexte von SHK1 somit nicht problemlos mit denen von SHK2 verglichen werden konnten²⁸⁷. Aus diesem Grund habe ich mich dazu entschlossen, zwischen SHK1 und SHK2

²⁸⁵ Hier sind alle Annotationskategorien außer *Pseudotreffer* gemeint, da unter den für die IAA-Berechnung ermittelten 100 *mit/für*-PP-Kontexten keine Pseudotreffer vorlagen.

²⁸⁶ Die Berechnung des κ -Werts erfolgte über die Website VassarStats von Richard Lowry (2001-2009) (<http://vassarstats.net/kappa.html>) (Mai 2019).

²⁸⁷ Beispielsweise wurde die *für*-PP im Ausdruck *Dank für X* zunächst als *attributive PP* annotiert, später wurde die Zuordnung dieser PP aufgrund des formelhaften Charakters dieser PP zur Kategorie *Formelhafte Sequenz* festgelegt. Die regulären Annotationen wurden stetig an die jeweils aktuelle Version des Annotationshandbuchs angepasst, die 100 *mit/für*-Annotationen von SHK1 enthielten jedoch noch die veraltete Zuordnung. Die Anfertigung aktualisierter bzw. neuer Annotationen durch

zumindest die Annotationen in den Annotationskategorien *Objekt-PP* und *adverbiale PP* zu vergleichen. Die Annotationen der übrigen Annotationskategorien habe ich als ‚Sonstige‘ zusammengefasst und ebenfalls berücksichtigt. Die Berechnung des entsprechenden IAA zwischen SHK1 und SHK2 lieferte einen Wert von $\kappa = 0,63$. Dieser Wert ist niedriger als die Übereinstimmung zwischen SHK2 und mir (in allen Annotationskategorien), liegt jedoch noch im akzeptablen Bereich²⁸⁸.

Im weiteren Annotationsprozess wurden die PP-Kontexte mit *mit/für* unter SHK1, SHK2 und mir aufgeteilt und entsprechend den Annotationskategorien (s. Kap. 10.4.1) und unter Rückgriff auf das Annotationshandbuch in mehreren Phasen annotiert. In wöchentlich stattfindenden Besprechungen zwischen der jeweiligen Hilfskraft und mir wurden Unsicherheiten und offene Fragen diskutiert und das Handbuch bei Bedarf überarbeitet und die Annotationen nochmals überprüft und ggf. korrigiert. In einem separaten Dokument wurden Annotationsbeispiele festgehalten (s. Annotationshandbuch im Anhang). Den Hilfskräften wurde empfohlen, dieses Dokument während der Annotationsphase zu konsultieren. Das Dokument hat sich als vorteilhaft erwiesen, da es zu einer konsistenten Annotation beitrug. Am Ende des Annotationsprozesses wurden die Annotationen nochmals sichtprobenartig überprüft und im Falle von Abweichungen zu den Annotationsrichtlinien im Annotationshandbuch korrigiert. Die Korrekturen sind jeweils dokumentiert.

10.5 Datenauswertung

Im Rahmen der vorliegenden Fallstudie wurden sowohl neue Daten im Lernerkorpus MERLIN erhoben und aufbereitet als auch auf eine Teilmenge bereits erhobener und aufbereiteter Daten aus Fallstudie 2 (Kap. 9) zurückgegriffen. Aus Fallstudie 2 liegen mir aufbereitete PP-Kontexte mit den desemantisierten Präpositionen *an/auf* vor, von denen lediglich alle als Objekt-PP annotierte PP-Kontexte in die Datenauswertung der vorliegenden Fallstudie einfließen. Bei den semantisierten Präpositionen *mit/für* werden für die Datenauswertung ebenfalls lediglich PP-Kontexte ausgewertet, die als Objekt-PP annotiert wurden.

SHK1 war nicht möglich, da diese zum Zeitpunkt der entsprechenden IAA-Berechnung nicht mehr als SHK tätig war.

²⁸⁸ Insgesamt ist das IAA damit niedriger als in Fallstudie 2 ($\kappa=0.9$, s. Kap. 9.4.2), damit liegen in der vorliegenden Fallstudie weniger Übereinstimmungen zwischen den Annotatorinnen vor. Der Unterschied im κ -Wert spiegelt wider, dass die Zuordnung der *mit/für*-PP-Kontexte zu den jeweiligen Annotationskategorien, und vor allem zu den syntaktischen Funktionen Objekt, Adverbiale und Attribut, den Annotatorinnen insgesamt schwerer fiel als in Fallstudie 2, in der *an/auf*-PP-Kontexte im Fokus standen. Die Schwierigkeit wurde auch während des Annotationsprozesses seitens der Annotatorinnen geäußert. Meiner Ansicht nach zeigen die unterschiedlichen κ -Werte indirekt an, dass die *mit/für*-PP aufgrund des ähnlichen semantischen Gehalts der Präpositionen in verschiedenen syntaktischen Funktionen der PP eine eindeutige Zuordnung zu einer bestimmten syntaktischen Funktion erschweren, während die PP mit *an/auf* bereits auf Basis ihres distinktiven semantischen Gehalts in unterschiedlichen syntaktischen Funktionen eindeutiger zugeordnet werden können.

Zunächst werden bei Objekt-PP mit *mit/für* alle PP-Kontexte aussortiert, die sich ebenfalls in den den Lernertexten zugrunde liegenden Aufgabenstellungen²⁸⁹ finden. Das Vorgehen entspricht dem von Fallstudie 2 und ist an entsprechender Stelle bereits ausführlich beschrieben worden (s. Kap. 9.5). An dieser Stelle werde ich den Vorgang daher nur kurz skizzieren. In den zugrundeliegenden Aufgabenstellungen werden zunächst alle Verben bzw. Adjektive ermittelt, die eine Präpositionsreaktion *mit* bzw. *für* aufweisen und damit Objekt-PP mit *mit/für* realisieren. Im Anschluss werden diese Verben/Adjektive mit denjenigen Verben/Adjektiven abgeglichen, die von den Lernenden produziert und damit im Rahmen der Datenaufbereitung der vorliegenden Fallstudie in den entsprechenden Excel-Dokumenten erfasst wurden (s. Kap. 9.4.1). Bei vorliegender Übereinstimmung werden die entsprechenden PP-Kontexte aus der weiteren Datenauswertung ausgeschlossen. Dabei werden lediglich die PP-Kontexte aussortiert²⁹⁰, die eindeutig zur entsprechenden Aufgabenstellung (die die ermittelten Verben/Adjektive mit entsprechender Präpositionsreaktion enthält) zugeordnet werden können.

Die folgende Tabelle (Tab. 62) gibt einen Überblick über die Objekt-PP-Kontexte mit den semantisierten Präpositionen *mit/für* sowie den desemantisierten Präpositionen *an/auf* je GeR-Gesamtniveau, die im Anschluss an die Datenaufbereitung und -bereinigung in die weitere Datenauswertung miteinfließen. Beim GeR-Gesamtniveau sind die Kriterien- und die Plus-Niveaus zusammengefasst (A2 repräsentiert A2 und A2+). Im Folgenden fasse ich die Präposition *mit/für* als semantisierte Präpositionen und *an/auf* als desemantisierte Präpositionen jeweils zusammen; die Abkürzung *semP* steht für semantisierte Präpositionen, die Abkürzung *desemP* für desemantisierte Präpositionen.

GeR-Gesamtniveau	Präposition						Summe
	semP			desemP			
	Gesamt	<i>mit</i>	<i>für</i>	Gesamt	<i>an</i>	<i>auf</i>	
A2	25	15	10	61	4	57	86
B1	122	52	70	131	24	107	253
B2	239	119	120	163	76	87	402
C1	49	15	34	24	15	9	73
Gesamt	435	201	234	379	119	260	814

Tabelle 62: PP-Kontexte je Präposition (*semP/desemP*) und GeR-Gesamtniveau (*bereinigt*).

Weiterhin wurde zusätzlich die Anzahl unterschiedlicher Lernertexte ermittelt, auf die sich die 814 PP-Kontexte insgesamt, sowie je Präposition (*semP/desemP*) und je GeR-Gesamtniveau, verteilen. Tabelle 63 illustriert die Anzahl unterschiedlicher Lernertexte, d. h. Lernender im Datensatz der vorliegenden Fallstudie.

²⁸⁹ Die einzelnen Aufgabenstellungen sind online auf der Korpus-Website verfügbar (https://merlin-platform.eu/C_mcorpus.php#anchor3, Mai 2019).

²⁹⁰ Entsprechende PP-Kontexte wurden aus pragmatischen Gründen in die Kategorie *Pseudotreffer* einsortiert. Dieses Vorgehen entspricht ebenfalls dem Vorgehen in Fallstudie 2 (s. Kap. 9.5).

Präposition	GeR-Gesamtniveau				
	A2	B1	B2	C1	Gesamt
Beide	71	179	209	30	489
semP	22	91	146	24	283
desemP	57	114	122	15	308

Tabelle 63: Anzahl unterschiedlicher Lernertexte je GeR-Gesamtniveau: Gesamtverteilung und getrennt nach semantischem Gehalt der Präposition.

Die insgesamt 814 PP-Kontexte von Objekt-PP mit semantisierten und desemantisierten Präpositionen verteilen sich auf insgesamt 489 Lernende (s. Tab. 63). Bei den semantisierten Präpositionen sind 283 unterschiedliche Lernertexte, d. h. Lernende repräsentiert, bei desemantisierten Präpositionen sind es 308 unterschiedliche Lernertexte, d. h. Lernende. Damit sind in der vorliegenden Fallstudie weitaus mehr Lernende repräsentiert als in Datensammlungen der traditionellen Z/FSE-Forschung, die i.d.R. weit weniger Lernende umfassen (vgl. Granger 2008: 269, Granger et al. 2015: 1).

Die im finalen Datensatz enthaltenen PP-Kontexte semantisierter und desemantisierter Präpositionen werden im Folgenden entsprechend der anfangs formulierten Leitfragen ausgewertet. Die Leitfragen seien an dieser Stelle nochmals genannt:

1. Welche Unterschiede zeigen sich bei der Korrektheit im Gebrauch semantisierter Präpositionen im Vergleich zu desemantisierten Präpositionen in Objekt-PP?
2. Welche Fehlertypen treten bei *semantisierten Präpositionen* häufig auf und (wie) unterscheiden sich diese von den Fehlertypen im Gebrauch von *desemantisierten Präpositionen* in Objekt-PP?

Für beide Leitfragen wird jeweils das unterschiedliche GeR-Gesamtniveau der Lernenden bei der Datenauswertung berücksichtigt.

10.5.1 Analysemethoden

Die ermittelten PP-Kontexte beider L1-Subkorpora werden mittels der computerunterstützten Fehleranalyse (computer-aided error analysis, CEA) (Dagneaux/Denness/Granger 1998) sowie der kontrastiven Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (Granger 1996, 2015) ausgewertet. Die nachfolgende Tabelle (Tab. 64) gibt einen Überblick über die in der vorliegenden Fallstudie angewandten Analysemethoden und die damit verbundenen Ziele. Nähere Informationen zur Anwendung dieser Methoden im Kontext der Dissertation finden sich in Kapitel 7.1.

Analysemethode	Ziel
Computerunterstützte Fehleranalyse	Ermittlung der Abweichungen (bzw. Korrektheit) im Präpositionsgebrauch (in zielsprachlich geforderten und zielsprachlich nicht geforderten PP-Kontexten) je semantischen Gehalt der Präposition (desemP und semP) und GeR-Gesamtniveau.
Kontrastive Interlanguage-Analyse	Kontrastiver Vergleich der Ergebnisse der computerunterstützten Fehleranalyse je GeR-Gesamtniveau und semantischem Gehalt der Präposition (desemP und semP).

Tabelle 64: Analysemethoden der Fallstudie zur Variable Semantischer Gehalt der Präposition: Überblick.

Im folgenden Kapitel werden die Ergebnisse der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse und ihre statistischen Auswertungen präsentiert.

10.5.2 Ergebnisse

Im vorliegenden Kapitel werden die Ergebnisse der quantitativen Datenauswertung vorgestellt, die eingangs formulierten Hypothesen überprüft und die entsprechenden Leitfragen beantwortet. Die Präsentation der Ergebnisse der statistischen Datenauswertung ist entsprechend der eingangs formulierten Leitfragen gegliedert: Zuerst gehe ich auf die Ergebnisse zur Korrektheit im Gebrauch von semP und desemP ein (Leitfrage 1). Im Anschluss daran werden die Ergebnisse der quantitativen Auswertung der Fehlertypen im Präpositionsgebrauch dargelegt (Leitfrage 2). Dabei wird jeweils überprüft, ob bzw. inwieweit der semantische Gehalt der Präposition einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch ausübt. In den Auswertungen werden jeweils die GeR-Gesamtniveaus der Lernenden berücksichtigt und miteinander kontrastiert.

10.5.2.1 Ergebnisse Auswertung Korrektheit im Präpositionsgebrauch

Die statistische Auswertung der computerunterstützten Fehleranalyse und der kontrastiven Interlanguage-Analyse liefert in Bezug auf die Korrektheit im Präpositionsgebrauch zwei zentrale Ergebnisse:

- 1) Insgesamt kann eine hohe Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch ermittelt werden.
- 2) Der semantische Gehalt von Präpositionen hat einen signifikanten Effekt auf die Korrektheit im Präpositionsgebrauch: Lernende zeigen bei semP eine deutlich höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch im Vergleich zu desemP. Dies zeigt sich unabhängig vom GeR-Gesamtniveau der Lernenden.
- 3) Das GeR-Gesamtniveau hat einen signifikanten Effekt auf die Korrektheit im Präpositionsgebrauch: Die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch nimmt mit steigendem GeR-Gesamtniveau stetig zu. Dies zeigt sich unabhängig vom semantischen Gehalt der Präposition.

Die folgende Tabelle (Tab. 65) gibt zunächst einen Überblick zur Korrektheit im Präpositionsgebrauch in Objekt-PP im vorliegenden Datensatz, getrennt nach semP (*mit/für*) und desemp (*an/auf*).

Präposition	PP-Kontexte (Gesamt)		Inkorrekte PP-Kontexte		Korrekte PP-Kontexte	
	semP	desemp	semP	desemp	semP	desemp
semP	435	100%	72	17%	363	83%
desemp	379	100%	123	32%	256	68%

Tabelle 65: Korrektheit im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemp): Überblick.

Tabelle 65 illustriert, dass unabhängig vom semantischen Gehalt der Präposition der korrekte Gebrauch von Präpositionen in Objekt-PP recht hoch ist (83% bzw. 68%). In Bezug auf mögliche Gebrauchsschwierigkeiten zeigen die Werte aber ebenfalls an, dass sich semP und desemp in Bezug auf die Korrektheit im Präpositionsgebrauch dennoch deutlich voneinander unterscheiden: Bei semP liegt fast ein doppelt so hoher Anteil an korrekten PP-Kontexten vor wie bei desemp (17% zu 32%).

Die nachfolgende Tabelle (Tab. 66) gibt einen Überblick zu korrekten und inkorrekten PP-Kontexten je semantischem Gehalt der Präposition und GeR-Gesamtniveau. Die Anzahl unterschiedlicher Lernertexte, d. h. Lernender, ist ebenfalls jeweils angegeben.

GeR-Gesamtniveau	Präposition					
	semP			desemp		
	Korrekte PP-Kontexte (%)	Inkorrekte PP-Kontexte (%)	AL	Korrekte PP-Kontexte (%)	Inkorrekte PP-Kontexte (%)	AL
A2	15 (60%)	10 (40%)	22	21 (34%)	40 (66%)	57
B1	94 (77%)	28 (23%)	91	91 (69%)	40 (31%)	114
B2	206 (86%)	33 (14%)	146	123 (75%)	40 (25%)	122
C1	48 (98%)	1 (2%)	24	21 (87%)	3 (13%)	15
Summe	363 (83%)	72 (17%)	283	256 (68%)	123 (32%)	308

Tabelle 66: Korrektheit im Präpositionsgebrauch in Objekt-PP Präposition und GeR-Gesamtniveau. AL=Anzahl Lernertexte.

Die Verteilung in Tabelle 66 spiegelt die Tendenzen der Gesamtergebnisse aus Tabelle 65 z.T. wider: Innerhalb von Objekt-PP zeigen semP (*mit/für*) durchweg eine höhere Korrektheit im Präpositionsgebrauch als desemp (*an/auf*). Gleichzeitig verdeutlichen die Werte in Tabelle 66, dass die Korrektheit im Präpositionsgebrauch mit zunehmendem GeR-Gesamtniveau stetig zunimmt, und zwar unabhängig vom semantischen Gehalt der Präposition.

Die vorliegenden Daten werden statistisch ausgewertet. Hierfür wähle ich die Regressionsanalyse und berechne gemischte Modelle (Näheres zur statistischen Auswertung s. Kap. 7.2). Die Korrektheit im Präpositionsgebrauch in Objekt-PP wird im Regressionsmodell als Zielvariable (abhängige Variable) festgelegt, der semantische Gehalt der Präposition (Praeposition) sowie das GeR-Gesamtniveau der Lernenden (Kompetenzniveau.op) werden als unabhängige Variablen berücksichtigt und stellen die festen Faktoren dar. Der/die einzelne Lernende wird jeweils als Zufallsfaktor („random effect“) ins Regressionsmodell integriert

(1|Lerner). Im ersten Regressionsmodell (RM1) wird ermittelt, ob eine Interaktion zwischen den festen Faktoren vorliegt, d. h. es wird überprüft, ob sich die Variable *Semantischer Gehalt der Präposition* und die Variable *GeR-Gesamtniveau* gegenseitig beeinflussen. Das entsprechende Modell (RM1) und sein Output sind im Folgenden angegeben:

RM1:

Korrektheit ~ Praeposition*Kompetenzniveau.op + (1|Lerner), data = dat, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.7150	0.3000	-2.383	0.017177
PraepositionsemP	1.1741	0.5427	2.164	0.030501
Kompetenzniveau.opB1	1.6109	0.3801	4.238	2.26e-05
Kompetenzniveau.opB2	1.9293	0.3798	5.080	3.78e-07
Kompetenzniveau.opC1	2.8271	0.7474	3.783	0.000155
PraepositionsemP:Kompetenzniveau.opB1	-0.7568	0.6212	-1.218	0.223063
PraepositionsemP:Kompetenzniveau.opB2	-0.4024	0.6031	-0.667	0.504630
PraepositionsemP:Kompetenzniveau.opC1	0.7919	1.3228	0.599	0.549394

Weitere statistische Berechnungen zeigen, dass keine allgemeine Interaktion zwischen den Variablen *Semantischer Gehalt der Präposition* und *GeR-Gesamtniveau* vorliegt ($p=0.4300444$). Das Ergebnis des Regressionsmodells RM1 ist in der folgenden Abbildung (Abb. 42) visualisiert. Die Abbildung zeigt auf der Y-Achse die Wahrscheinlichkeit²⁹¹ für die Zielvariable (hier: Korrektheit im Präpositionsgebrauch), auf der X-Achse sind die einzelnen GeR-Gesamtniveaus angegeben. Die durchgezogene Linie repräsentiert die ermittelten Wahrscheinlichkeiten des korrekten Präpositionsgebrauchs für desemP, die gestrichelte Linie zeigt die entsprechenden Wahrscheinlichkeiten für semP. Für jedes GeR-Gesamtniveau ist der Mittelwert angegeben (Kreis) sowie die Standardabweichung zum jeweiligen Mittelwert (senkrechte Linie mit waagerechten Linien).

²⁹¹ Beispiel: Der Wert 0.6 repräsentiert 60%.

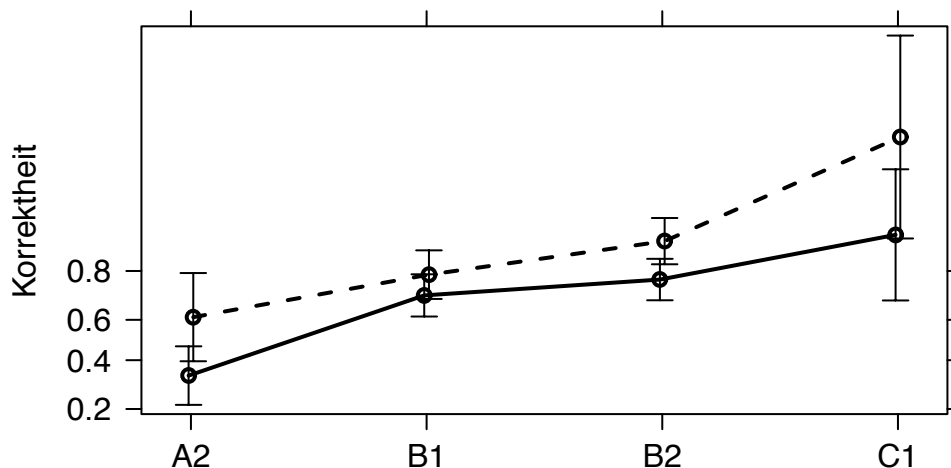


Abbildung 42: Korrektheit im Präpositionsgebrauch Präposition (semP/desemP) und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM1, desemP = durchgezogene Linie, semP = gestrichelte Linie.

Abbildung 42 zeigt, dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch innerhalb von Objekt-PP für semP (*mit/für*) in jedem GeR-Gesamtniveau höher liegt als für desemP (*an/auf*). Deutliche Unterschiede sind vor allem auf dem Niveau elementarer Sprachverwendung (A2) sowie dem höheren Niveau selbstständiger Sprachverwendung (B2) festzustellen – hier überlappen die jeweiligen Standardabweichungen jeweils nicht die entsprechenden Mittelwerte. Dahingehen sind die Unterschiede in der Wahrscheinlichkeit für Korrektheit geringer im niedrigen Niveau selbstständiger Sprachverwendung (B1) sowie auf dem Niveau der kompetenten Sprachverwendung (C1) (vgl. Europarat 2001: 41) – hier überlappen die Standardabweichungen die jeweiligen Mittelwerte, d. h. die Unterschiede in den Wahrscheinlichkeiten für Korrektheit im Präpositionsbereich bei semP und desemP sind auf diesen Niveaus statistisch gesehen nicht relevant. Gleichzeitig illustriert die Visualisierung (Abb. 42), dass die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch mit steigendem GeR-Gesamtniveau zunimmt, und zwar sowohl für desemP als auch für semP.

Wie oben erwähnt, konnte in RM1 keine Interaktion zwischen den Variablen *Semantischer Gehalt der Präposition* und *GeR-Gesamtniveau* ermittelt werden. Das bedeutet, dass der semantische Gehalt der Präposition und das GeR-Gesamtniveau sich im Hinblick auf die Korrektheit im Präpositionsgebrauch nicht gegenseitig beeinflussen. Da keine Interaktion zwischen den beiden Variablen vorliegt, wird das Regressionsmodell RM1 modifiziert und für das modifizierte Modell lediglich die Einzeleffekte der Variablen berücksichtigt. Das modifizierte Modell RM1a ist im Folgenden mit seinem Output angegeben. Das Ergebnis des Modells RM1a ist in Abbildung 43 visualisiert.

RM1a:

```
Korrektheit ~ Praeposition + Kompetenzniveau.op + (1|Lerner), data = dat, family = "binomial", control = glmerControl(optimizer = "bobyqa")
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.5816	0.2550	-2.281	0.022571
PraepositionsemP	0.7258	0.1956	3.710	0.000207
Kompetenzniveau.opB1	1.3450	0.3074	4.376	1.21e-05
Kompetenzniveau.opB2	1.8166	0.3112	5.838	5.29e-09
Kompetenzniveau.opC1	3.1894	0.6230	5.120	3.06e-07

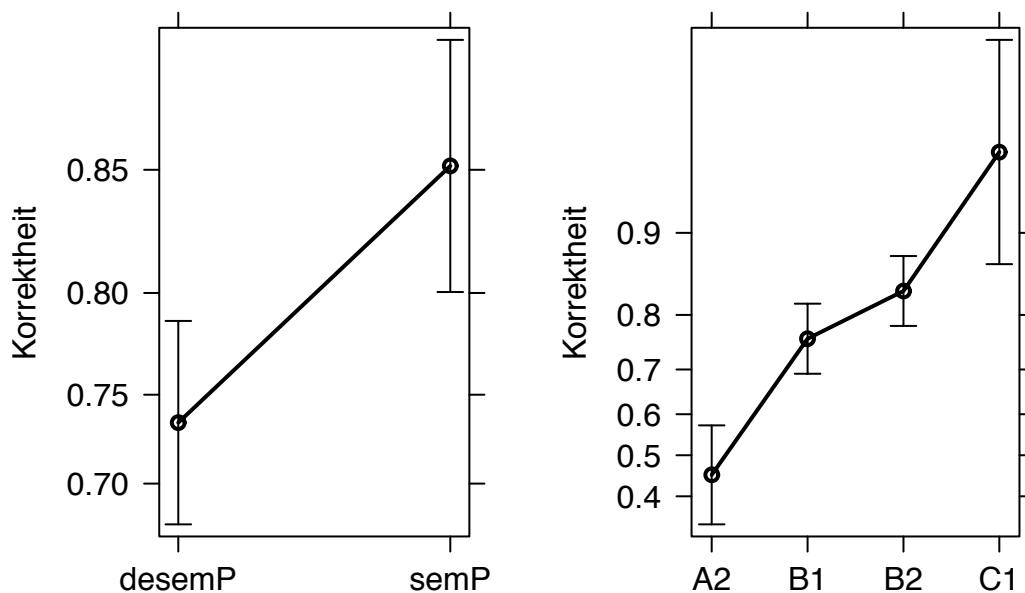


Abbildung 43: Korrektheit im Präpositionsgebrauch in Objekt-PP Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM1a.

Die Werte im Modell-Output (s.o.) zeigen, dass sowohl für den semantischen Gehalt der Präposition als auch für die jeweiligen GeR-Gesamtniveaus jeweils signifikante Einzeleffekte auf die Korrektheit im Präpositionsgebrauch innerhalb von Objekt-PP ermittelt werden können. Ein signifikanter allgemeiner Effekt für das GeR-Gesamtniveau liegt ebenfalls vor, wie weitere Berechnungen zeigen ($p=2.966e-09$). Abbildung 43 visualisiert, dass semP (*mit/für*) innerhalb von Objekt-PP eine signifikant höhere Wahrscheinlichkeit für den korrekten Präpositionsgebrauch aufweisen als desemP (*an/auf*) (ca. 74% zu 85%), und das unabhängig vom GeR-Gesamtniveau der Lernenden (s. Abb. 43). Gleichzeitig greift ein Einzeleffekt des GeR-Niveaus: Die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch in Objekt-PP nimmt mit steigendem GeR-Gesamtniveau zu, und das unabhängig vom semantischen Gehalt der Präposition. Nimmt man an, dass die Korrektheit im Präpositionsgebrauch den Erwerb von Präpositionen und PP sowie den Grad ihrer Beherrschung widerspiegelt, so lässt sich an dieser Stelle ableiten, dass DaF-Lernende den Gebrauch von Präpositionen in Objekt-PP mit steigendem Kompetenzniveau immer besser beherrschen, d. h. sich dem zielsprachlichen Gebrauch annähern, jedoch im Gebrauch von semP (*mit/für*) insgesamt eine höhere Sicherheit zeigen als im Gebrauch von desemP (*an/auf*).

Die Evaluation des modifizierten Regressionsmodells RM1a zeigt, dass das Modell die Datenverteilung im vorliegenden Datensatz zu ca. 17,6% erklären kann ($R^2=0.1758991$). Dieser Wert ist im Rahmen der vorliegenden Fallstudie als relativ hoch einzustufen. Das

Regressionsmodell modelliert die Korrektheit im Präpositionsgebrauch innerhalb von Objekt-PP und überprüft, inwieweit der semantische Gehalt der Präposition und das GeR-Gesamtniveau einen Effekt auf diese ausübt. Geht man davon aus, dass der Sprachgebrauch von Lernenden und damit auch der korrekte Gebrauch von Präpositionen in Objekt-PP durch eine Vielzahl von Faktoren beeinflusst wird, so deutet die Evaluation des Modells RM1a (R^2 -Wert) darauf hin, dass der semantische Gehalt der Präposition sowie das GeR-Gesamtniveau der Lernenden ein nicht zu vernachlässigendes Erklärungspotenzial für die Korrektheit im Präpositionsgebrauch bieten.

Zusammenfassung

Die erste Leitfrage kann den Ergebnissen der statistischen Auswertung nach wie folgt beantwortet werden:

1. Welche Unterschiede zeigen sich bei der Korrektheit im Gebrauch semantisierter Präpositionen im Vergleich zu desemantisierten Präpositionen in Objekt-PP?

Insgesamt zeigt sich in Objekt-PP sowohl bei semP als auch bei desemP eine hohe Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch. Die Auswertung zeigt jedoch, dass semP (*mit/für*) eine signifikant höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch zeigen als desemP (*an/auf*): Es kann ein signifikanter Effekt des semantischen Gehalts der Präposition auf die Korrektheit im Präpositionsgebrauch ermittelt werden. Dieses Ergebnis gilt unabhängig vom GeR-Gesamtniveau der DaF-Lernenden. Weiterhin zeigen die statistischen Auswertungen, dass parallel dazu ebenfalls das GeR-Gesamtniveau – unabhängig vom semantischen Gehalt der Präposition in Objekt-PP – einen signifikanten Effekt auf die Wahrscheinlichkeit für die Korrektheit im Präpositionsgebrauch zeigt.

Die anfangs aufgestellte 1. Hypothese kann somit bestätigt werden: Innerhalb von Objekt-PP zeigt sich für semP (*mit/für*) im Vergleich zu desemP (*an/auf*) eine signifikant höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch. Es kann ein entsprechender signifikanter Effekt ermittelt werden. Dieser Effekt könnte in der zuverlässigen Form-Funktionsverbindung von semP (*mit/für*) in Objekt-PP begründet sein. Geht man davon aus, dass zuverlässige Form-Funktionsverbindungen positiv im L2-Erwerb wirken, könnte die signifikant höhere Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch bei semP diesen Vorteil im Erwerb widerspiegeln. Weiterhin bestätigen die statistischen Auswertungen ebenfalls die 2. Hypothese, dass mit steigendem GeR-Gesamtniveau die Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch unabhängig vom semantischen Gehalt der Präposition zunimmt: Auch hier konnte ein entsprechender allgemeiner signifikanter Effekt ermittelt werden. Somit zeigen die Ergebnisse, dass DaF-Lernende ihre Kompetenz im Gebrauch von semP (*mit/für*) und desemP (*an/auf*) in Objekt-PP mit zunehmenden GeR-Gesamtniveau ausbauen, wobei sich dennoch größere Unsicherheiten im Gebrauch von desemP prognostizieren lassen.

10.5.2.2 Ergebnisse Auswertung Fehlertypen im Präpositionsgebrauch

Für die Auswertung der Fehlertypen im Präpositionsgebrauch werden ausschließlich die inkorrekten PP-Kontexte von semP und desemp berücksichtigt. Die statistischen Auswertungen der computerunterstützten Fehleranalyse sowie der kontrastiven Interlanguage-Analyse liefern in Bezug auf die Fehlertypen im Präpositionsgebrauch folgende zentrale Ergebnisse:

- 1) Der semantische Gehalt der Präposition hat einen signifikanten Einfluss auf den Fehlertyp im Präpositionsgebrauch in Objekt-PP: Die Wahrscheinlichkeit für Auslassungen von Präpositionen und wahlbedingte Fehler im Präpositionsgebrauch ist signifikant höher für desemp; die Wahrscheinlichkeit für inkorrekte Hinzufügungen von Präpositionen ist dahingegen signifikant höher für semP.
- 2) Das GeR-Gesamtniveau der Lernenden zeigt einen signifikanten Effekt auf Auslassungen von Präpositionen und wahlbedingte Präpositionsfehler: Die Wahrscheinlichkeit für Auslassungen nimmt mit steigendem GeR-Gesamtniveau ab, die Wahrscheinlichkeit für wahlbedingte Fehler nimmt dahingegen zu. Für Hinzufügungen von Präpositionen kann kein Effekt des GeR-Gesamtniveaus ermittelt werden.

Die Verteilung korrekter und inkorrekt PP-Kontexte je semantischem Gehalt der Präposition und GeR-Gesamtniveau habe ich bereits in Tabelle 66 präsentiert. An dieser Stelle (s. Tab. 67) gebe ich die absolute Verteilung inkorrekt PP-Kontexte je semantischen Gehalt der Präposition (desemp und semP) und GeR-Gesamtniveau im vorliegenden Datensatz erneut an. Die Anzahl unterschiedlicher Lernertexte, d. h. Lernenden, die im Datensatz inkorrekt PP-Kontexte repräsentiert sind, ist ebenfalls angegeben.

GeR-Gesamtniveau	Präposition		AL Gesamt
	semP	desemp	
	Inkorrekte PP-Kontexte	Inkorrekte PP-Kontexte	
A2	10	40	46
B1	28	40	61
B2	33	40	62
C1	1	3	3
Summe	72	123	172

Tabelle 67: Inkorrekte PP-Kontexte je Präposition (semP/desemp) in Objekt-PP und GeR-Gesamtniveau. AL= Anzahl Lernertexte.

Wie man den Werten in Tabelle 67 entnehmen kann, sind im Datensatz inkorrekt PP-Kontexte mit semP und desemp insgesamt 172 verschiedene DaF-Lernende repräsentiert. Damit liegen keine eindeutigen Ausreißer vor, d. h. die inkorrekten PP-Kontexte wurden nicht einigen wenigen DaF-Lernenden produziert.

Für die Auswertung der inkorrekten PP-Kontexte mit semP und desemp werden die im MERLIN vorliegenden (korrigierten) und manuell hinzugefügten Fehlerannotationen im Bereich Präpositionsgebrauch ausgewertet (s. Kap. 10.4.1). Die folgende Tabelle (Tab. 68) gibt

einen ersten Überblick zur Verteilung einzelner Fehlertypen im Präpositionsgebrauch in Objekt-PP je semantischem Gehalt der Präposition.

Präposition	Fehlertyp im Präpositionsgebrauch				
	Wahl	Auslassung	Position	Hinzufügung	Summe
semP	12 (17%)	9 (13%)	0 (%)	51 (71%)	72 (100%)
desemP	40 (33%)	63 (51%)	1 (1%)	19 (15%)	123 (100%)

Tabelle 68: Fehlertypen in Objekt-PP je Präposition (semP/desemP) (GeR-Gesamtniveaus A2, B1, B2 und C1): Überblick.

Im Gesamtüberblick zu den Fehlertypen im Präpositionsgebrauch (Tab. 68) fällt zunächst auf, dass je semantischem Gehalt der Präposition in Objekt-PP unterschiedliche Fehlertypen dominieren: Im Falle von semP (*mit/für*) dominiert der Fehlertyp *Hinzufügung*, d. h. die Realisierung von zielsprachlich nicht geforderten Präpositionen bzw. PP: Über 70% aller inkorrekten PP-Kontexte sind diesem Fehlertyp zugeordnet. Bei desemP (*an/auf*) dominiert dahingehen der Fehlertyp *Auslassung*, d. h. Präpositionen werden in zielsprachlichen Kontexten nicht von den Lernenden realisiert. Hier repräsentiert ca. die Hälfte aller inkorrekten PP-Kontexte (51%) diesen Fehlertyp. Der Fehlertyp *Wahl*, d. h. der Gebrauch zielsprachlich inkorrektur Präpositionen, ist sowohl bei semP als auch bei desemP der zweithäufigste Fehlertyp (17% bei semP bzw. 33% bei desemP). Unabhängig vom semantischen Gehalt ist der Fehlertyp *Position* (fast) nicht vertreten (0% bei semP bzw. 1% bei desemP) und kann daher für die Auswertungen vernachlässigt werden. Die folgenden Beispiele aus dem Datensatz der vorliegenden Fallstudie illustrieren die häufigen Fehlertypen *Wahl*, *Auslassung* und *Hinzufügung* bei semP (*mit/für*) und desemP (*an/auf*). Diese Fehlertypen stehen im Fokus der folgenden Auswertung.

Fehlertyp Wahl:

1. *Aber bevor man sich auf etwas entscheidet oder irgendwelchen Objekt auswählt, [...]* (ID: 1031_0003388)
ZH1: Aber bevor man sich für etwas entscheidet oder irgendein Objekt auswählt, [...]
2. *[...] möchte ich noch hervorgeben, dass man ohne Kulturkenntnisse nicht in Gastland umgehen kann.* (ID: 1031_0002085)
ZH1: [...] möchte ich noch hervorgeben, dass man ohne Kulturkenntnisse nicht mit dem Gastland umgehen kann.
3. *Dann denke ich um eine kleine Papagei, [...]* (ID: 1061_0120859)
ZH1: Dann denke ich an einen kleinen Papagei, [...]
4. *Kann ich auch für Freizeitaktivitäten hoffen?* (ID: 1023_0107244)
ZH1: Kann ich auch auf Freizeitaktivitäten hoffen?

Fehlertyp Auslassung:

1. *Ich interessiere mich sehr das Wohnung* (ID: 1091_0000125)
ZH1: Ich interessiere mich sehr für die Wohnung.
2. *Ich habe Kleine Geschwezter die ich sehr gut umgehen kann, [...]* (ID: 1023_0109946)
ZH1: Ich habe kleine Geschwister , mit denen ich sehr gut umgehen kann, [...]
3. *Deshalb muss ich mein Kind aufpassen.* (ID: 1061_0120456)
ZH1: Deshalb muss ich auf mein Kind aufpassen.
4. *In Freizeit möchte ich viele Aktivitäten teilnehmen* (ID: 1023_0109914)
ZH1: In der Freizeit möchte ich an vielen Aktivitäten teilnehmen.

Fehlertyp Hinzufügung:

1. *ich gebe der Karte für dich am Abend, [...]* (ID: 1091_0000257)
ZH1: *gebe ich dir die Karte am Abend, [...]*
2. *Ich mag auch natürlich mit anderen Gebräuchen kennen lernen, [...]* (ID: 1031_0003211)
ZH1: Ich mag auch natürlich andere Gebräuche kennenlernen, [...]
3. *[...] leichter, auf Problemen beim Wohen oder Haussuche zu vermeiden* (ID: 1031_0003354)
ZH1: [...] leichter, Probleme beim Wohnen oder der Haussuche zu vermeiden
4. *Besonderes muss man an verschieden Agenturen beachten, [...]* (ID: 1031_0003071)
ZH1: Besonders muss man verschieden Agenturen beachten, [...] ²⁹²

Folgende Tabelle (Tab. 69) gibt die Verteilung der Fehlertypen Wahl, Auslassung und Hinzufügung je semantischem Gehalt der Präposition und GeR-Gesamtniveau wieder.

²⁹² Hier liegt auf Ebene der ZH1 ein Fehler vor; die Adjektivflexion (*verschieden*) wurde nicht korrigiert. Dies ist für die Erfassung des Fehlers im Präpositionsgebrauch jedoch nicht relevant.

GeR- Gesamtniveau	Präposition					
	semP			desemP		
	Wahl	Auslassung	Hinzufügung	Wahl	Auslassung	Hinzufügung
A2	1	0	9	8	28	3
B1	3	5	20	9	25	6
B2	8	4	21	22	10	8
C1	0	0	1	1	0	2
Summe	12	9	51	40	63	19

Tabelle 69: Fehlertypen Wahl, Auslassung und Hinzufügung in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau.

Da im GeR-Gesamtniveau C1 sehr wenige inkorrekte PP-Kontexte vorliegen, wird dieses GeR-Gesamtniveau in den folgenden statistischen Auswertungen nicht weiter berücksichtigt. Zusätzlich wird für die Auswertung der Fehlertypen *Wahl* und *Auslassung* das GeR-Gesamtniveau A2 nicht beachtet, da in den inkorrekten PP-Kontexten von semP keine Auslassungen und lediglich ein einziger wahlbedingter Fehler auf diesem Kompetenzniveau vorliegen. Im Folgenden werden zunächst die jeweils dominanten Fehlertypen Hinzufügung und Auslassung je semantischem Gehalt der Präposition und GeR-Gesamtniveau einzeln ausgewertet. Hier werden insgesamt 191 PP-Kontexte (Fehlertyp Hinzufügung) von 169 DaF-Lernenden bzw. 141 PP-Kontexte (Fehlertyp Auslassung) von 123 DaF-Lernenden ausgewertet (vgl. Tab. 67). Im Anschluss stelle ich die Auswertung des Fehlertyps Wahl vor, in der insgesamt ebenfalls 141 PP-Kontexte von 123 DaF-Lernenden berücksichtigt werden (vgl. Tab. 67).

Die Daten werden je Fehlertyp mittels einer Regressionsanalyse ausgewertet, wobei je Fehlertyp gemischte Modelle berechnet werden (Näheres zu den statistischen Auswertungen in der vorliegenden Dissertation s. Kap. 7.2). Als Zielvariable wird der jeweilige Fehlertyp festgelegt (Wahl, Auslassung, Hinzufügung). Der *semantische Gehalt der Präposition* (Praeposition) und das *GeR-Gesamtniveau* (Kompetenzniveau.op) werden als unabhängige Variablen berücksichtigt und stellen die festen Faktoren im jeweiligen Regressionsmodell dar. Äquivalent zu den Regressionsmodellen zur Korrektheit im Präpositionsgebrauch werden in den Regressionsmodellen zu den Fehlertypen der/die einzelne Lernende jeweils als ‚random effect‘, d. h. als Zufallsfaktor (1|Lerner) festgelegt. Im Modell werden ausschließlich inkorrekte PP-Kontexte berücksichtigt (in den Regressionsmodellen wird auf diese jeweils mit *dat.subset* referiert). Mittels der Regressionsanalyse wird je Fehlertyp überprüft, ob auf Basis des vorliegenden Datensatzes ein Effekt des Semantischen Gehalts der Präposition und/oder des GeR-Gesamtniveaus auf den jeweiligen Fehlertyp im Präpositionsgebrauch in Objekt-PP ermittelt werden kann. Im Folgenden werden die einzelnen Regressionsmodelle und ihre Ergebnisse einzeln präsentiert.

Das Regressionsmodell für den Fehlertyp Hinzufügung (RM2) ist unten mit seinem Output angegeben, das Ergebnis ist in Abbildung 44 visualisiert. Die Wahrscheinlichkeiten für den Fehlertyp sind auf der Y-Achse, die GeR-Gesamtniveaus auf der X-Achse angegeben. Die gestrichelte Linie repräsentiert die entsprechenden Wahrscheinlichkeiten für semP, die durchgezogene Linie repräsentiert die Wahrscheinlichkeiten für desemP.

RM2:

Hinzufuegung ~ Praeposition*Kompetenzniveau.op + (1|Lerner), data = dat.subset²⁹³, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.5123	0.6003	-4.185	2.85e-05
PraepositionsemP	4.7095	1.2131	3.882	0.000103
Kompetenzniveau.opB1	0.7777	0.7460	1.043	0.297149
Kompetenzniveau.opB2	1.1260	0.7188	1.567	0.117207
PraepositionsemP:Kompetenzniveau.opB1	-2.0586	1.3574	-1.517	0.129376
PraepositionsemP:Kompetenzniveau.opB2	-2.7636	1.3262	-2.084	0.037169

Weitere statistische Berechnungen zeigen, dass für die beiden Variablen *Semantischer Gehalt* der *Präposition* und *GeR-Gesamtniveau* keine allgemeine signifikante Interaktion vorliegt ($p=0.1082$), d. h. die Variablen beeinflussen sich nicht gegenseitig.

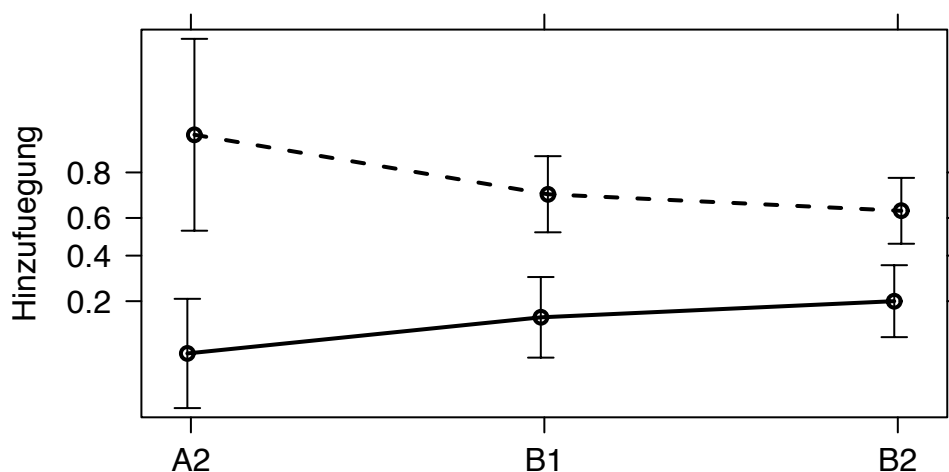


Abbildung 44: Fehlertyp Hinzufuegung im Praepositionsgebrauch in Objekt-PP je Praeposition (semP/desempP) und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM2, desempP = durchgezogene Linie, semP = gestrichelte Linie.

Wie man Abbildung 44 entnehmen kann, ist die Wahrscheinlichkeit für den Fehlertyp Hinzufuegung bei semP (*mit/für*) konstant höher als bei desempP (*an/auf*), und das unabhängig vom GeR-Gesamtniveau. Auf jedem GeR-Gesamtniveau zeigt sich jeweils ein deutlicher Unterschied in der Wahrscheinlichkeit des Fehlertyps je semantischem Gehalt der Präposition: Die Markierungen der Standardabweichungen überlappen nie die jeweiligen Mittelwerte (s. Abb. 44). Dies zeigt statistisch signifikante Unterschiede je GeR-Gesamtniveau an. Weiterhin kann beobachtet werden, dass für semP die Wahrscheinlichkeit für den Fehlertyp Hinzufuegung leicht abnimmt (über 80% - ca. 70% - ca. 60%) während sie für desempP leicht zunimmt (ca. 10% - ca. 18% - ca. 20%). Der größte Unterschied in der Wahrscheinlichkeit

²⁹³ *dat.subset* referiert auf die Gesamtmenge aller inkorrekten Objekt-PP-Kontexte mit semP und desempP.

für Hinzufügungen von Präpositionen zeigt sich dabei auf dem Niveau A2 – hier zeigen semP eine Wahrscheinlichkeit von über 80%, dessemP dahingegen lediglich ca. 10%. Mit steigendem GeR-Gesamtniveau nähern sich die Wahrscheinlichkeiten für den Fehlertyp Hinzufügung bei semP und dessemP an, unterscheiden sich dennoch auf dem höheren Niveau selbstständiger Sprachverwendung (B2) noch deutlich voneinander.

Da die Ergebnisse des Modells RM1 keine allgemeine Interaktion zwischen den beiden Variablen *Semantischer Gehalt der Präposition* und *GeR-Gesamtniveau* zeigen, werden in einem modifizierten Modell (RM2a) lediglich die Einzeleffekte dieser Variablen berücksichtigt. Das modifizierte Modell RM2a ist im Folgenden mit seinem Output angegeben. Abbildung 45 visualisiert das Ergebnis des entsprechenden Modells.

RM2a:

Hinzufuegung ~ Praeposition + Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.8980	0.4048	-4.688	2.76e-06
PraepositionsemP	2.6457	0.3767	7.023	2.17e-12
Kompetenzniveau.opB1	0.1662	0.4971	0.334	0.738
Kompetenzniveau.opB2	0.1181	0.4929	0.240	0.811

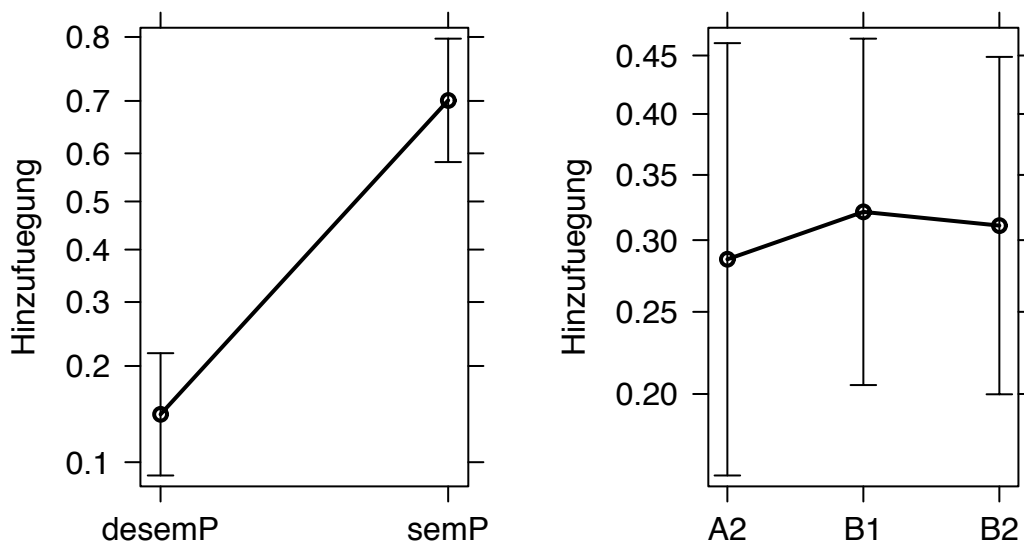


Abbildung 45: Fehlertyp Hinzufügung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM2a.

Dem Modell-Output des Regressionsmodells RM2a nach liegt kein Effekt des GeR-Gesamtniveaus im Vergleich zum Referenzniveau A2 vor, weder für B1 noch für B2 (s. p-Werte im Modell-Output oben). Abbildung 45 kann man entnehmen, dass die Wahrscheinlichkeit für den Fehlertyp Hinzufügung in den GeR-Gesamtniveaus A2, B1 und B2 sehr ähnlich ist – die Wahrscheinlichkeit schwankt zwischen ca. 29% und 31%, wobei in jedem untersuchten GeR-Gesamtniveau eine große Varianz vorliegt (s. Markierungen der Standardabweichungen der jeweiligen Mittelwerte in Abb. 45). Dahingehen lässt sich ein signifikanter Unterschied in der

Wahrscheinlichkeit für den Fehlertyp Hinzufügung je semantischem Gehalt der Präposition feststellen (s. p-Wert im Modell-Output und Abb. 45). Für *desemP* liegt die Wahrscheinlichkeit bei ca. 15%, für *semP* dahingegen bei 70% – das ist mehr als viermal so viel. Die Ergebnisse zeigen demnach, dass die Wahrscheinlichkeit für den Fehlertyp Hinzufügung in Objekt-PP vom semantischen Gehalt der Präposition beeinflusst wird. Das GeR-Gesamtniveau scheint den Berechnungen nach hier keinen (maßgebenden) Einfluss zu haben.

Das Regressionsmodell RM2a wird erneut angepasst, sodass im modifizierten Modell (RM2b) lediglich der Einzeleffekt des semantischen Gehalts der Präposition berücksichtigt wird. Das modifizierte Modell und der Output des Modells sind im Folgenden angegeben. Der Vollständigkeit halber ist das Ergebnis des modifizierten Modells (RM2b) in Abbildung 46 visualisiert.

RM2b:

Hinzufuegung ~ Praeposition + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.8015	0.2618	-6.882	5.92e-12
PraepositionsemP	2.6690	0.3690	7.233	4.71e-13

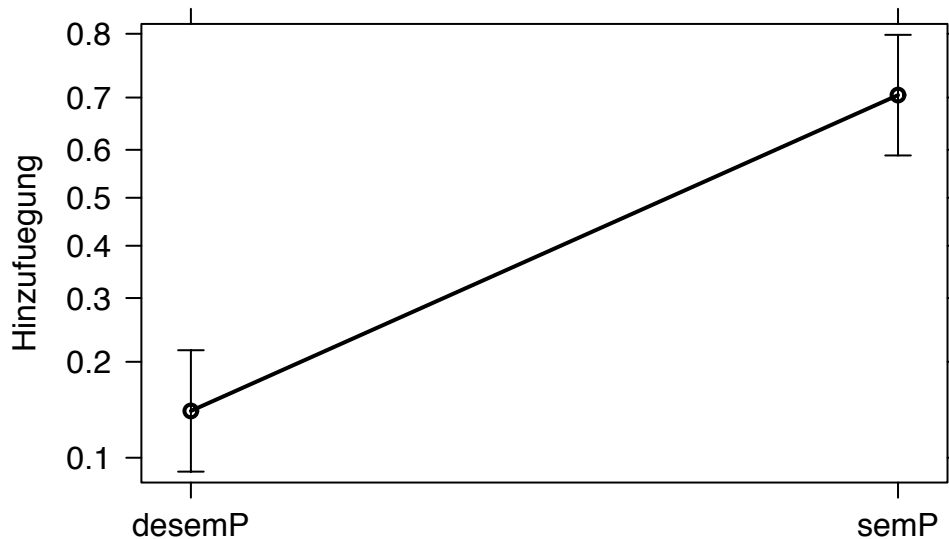


Abbildung 46: Fehlertyp Hinzufügung im Präpositionsgebrauch in Objekt-PP Präposition (*semP*/*desemP*) (Einzeleffekt). Ergebnis des Regressionsmodells RM2b.

Das Ergebnis des Modells RM2b (Werte im Modell-Output und Abb. 46) spiegelt den Einzeleffekt des semantischen Gehalts der Präposition auf den Fehlertyp Hinzufügung in Objekt-PP deutlich wider. Dieser Effekt gilt unabhängig vom Kompetenzniveau der Lernenden.

Die Evaluation des finalen Regressionsmodells (RM2b) zeigt, dass das Modell die Varianz im vorliegenden Datensatz zu 27,6% erklären kann ($R^2 = 0.2755648$). Das bedeutet, dass der semantische Gehalt der Präposition in Objekt-PP ein vergleichsweise relativ hohes Erklärungspotenzial für zielsprachlich inkorrekte Hinzufügungen von Präpositionen birgt, und

zwar dergestalt, dass semP eine signifikant höhere Wahrscheinlichkeit für Hinzufügungen zeigen als desemP. Die Hinzufügungen könnten durch die unterschiedliche Zuverlässigkeit der jeweiligen Form-Funktionsverbindung bzw. durch eine Übergeneralisierung der zuverlässigen Form-Funktionsverbindungen bei semP bedingt sein. Ich möchte dies kurz ausführen.

Lerneräußerungen wie die unteren lassen vermuten, dass Hinzufügungen von semP (vor allem die Präposition *für*) durch eine Übergeneralisierung der zuverlässigen Form-Funktionsverbindungen der entsprechenden PP erklärt werden können: SemP werden in PP-Kontexten hinzugefügt, in denen zielsprachlich eine Nominalphrase (NP) im Dativ gefordert ist, die auf den/die Nutznießende/n einer Handlung bzw. eine/n Rezipienten/-in verweist:

1. *also ich möchte eine Fahrkarte für dich schenken, [...]* (ID: 1091_0000257)
ZH1: also ich möchte dir eine Fahrkarte schenken, [...]
2. *Ich würde für jede Leser empfehlen neue Kulturen kennenzulernen.* (ID: 1031_0003240)
ZH1: Ich würde jedem Leser empfehlen neue Kulturen kennenzulernen.
3. *Um ein Uhr, haben wir das Geschenk für meiner Schwiegermutter gegeben.* (ID: 1061_0120478)
ZH1: Um ein Uhr haben wir meiner Schwiegermutter das Geschenk gegeben.
4. *[...], ich zeige unsere Traditionen, unsere Kultur und auch unsere Speisen für die Bürgern des Gastlandes.* (ID: 1031_0003231)
ZH1: [...], zeige ich unsere Traditionen, unsere Kultur und auch unsere Speisen den Bürgern des Gastlandes.

In Kapitel 1.2.3 habe ich darauf hingewiesen, dass eine *für*-PP (in- und außerhalb von Objekt-PP) auf den/die Nutznießende/n einer Handlung referieren kann (vgl. GDS 1997: 2131). Semantisch leistet ein Dativobjekt ähnliches (Duden 2016: 830). Dies spiegelt die semantische Nähe von *für*-PP und Dativ-NP wider und macht die Hinzufügungen der semP *für* in den obigen Beispielen semantisch nachvollziehbar. Vergangene Untersuchungen in der Z/FSE-Forschung haben ebenfalls festgestellt, dass Lernende in Kontexten von indirekten Objekten (hier: Dativobjekte) häufig eine Präposition hinzufügen, um diese Objekte auf diese Weise formal zu markieren (Pfaff 1984: 294, Kaltenbacher/Klages 2006: 87). Häufig eingesetzt werden hier PP mit den Präpositionen *zu*, *bei* und *für*, z. B. *der Pferd schenkt für eine Hund ein Keks* (entnommen aus Kaltenbacher/Klages 2006: 87); vermutlich geschieht dies aufgrund der semantischen Nähe entsprechender PP zum Dativobjekt und der vergleichsweise zuverlässigen Form-Funktionsverbindung, und damit konstanten semantischen Gehalts entsprechender Präpositionen bzw. PP²⁹⁴. Die Beispiele im vorliegenden Datensatz könnten somit anzeigen, dass DaF-Lernende die zuverlässigen Form-Funktionsverbindungen

²⁹⁴ Nach Breindl (1989: 39/42) können die Präpositionen *zu* und *bei* in Objekt-PP ebenfalls als semantisiert angesehen werden.

semantischer Präpositionen (hier vor allem *für*) registrieren und auf semantisch ähnliche Kontexte übergeneralisieren.

Für die Fehlertypen Auslassung und Wahl werden ebenfalls Regressionsmodelle (RM3 und RM4) berechnet, die im Aufbau den Regressionsmodellen zum Fehlertyp Hinzufügung entsprechen. Zur Erinnerung sei an dieser Stelle nochmals erwähnt, dass – anders als beim Fehlertyp Hinzufügung – für beide Fehlertypen (Auslassung und Wahl) das GeR-Gesamtniveau A2 nicht berücksichtigt wird, da im Datensatz für entsprechende Fehlertypen z.T. keine PP-Kontexte vorliegen (s. Tab. 69). Somit werden für die Fehlertypen lediglich die GeR-Gesamtniveaus B1 und B2 miteinander kontrastiert.

Zunächst möchte ich auf den Fehlertyp Auslassung eingehen. Das entsprechende Regressionsmodell (RM3), mit dem zunächst eine Interaktion zwischen der Variable *Semantischer Gehalt der Präposition* (Praeposition) und der Variable *GeR-Gesamtniveau* (Kompetenzniveau.op) unter der Berücksichtigung des/der einzelnen Lernenden als Zufallsfaktor (1|Lerner) überprüft wird, ist im Folgenden mit seinem Output angegeben²⁹⁵.

RM3:

Auslassung ~ Kompetenzniveau.op*Praeposition + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.5108	0.3266	1.564	0.117798
Kompetenzniveau.opB2	-1.6094	0.4899	-3.285	0.001019
PraepositionsemP	-2.0369	0.5917	-3.442	0.000577
Kompetenzniveau.opB2:PraepositionsemP	1.1545	0.8763	1.317	0.187697

Die Werte des Modell-Outputs (s.o.) zeigen, dass im Hinblick auf den Fehlertyp Auslassung keine signifikante Interaktion zwischen dem semantischen Gehalt der Präposition und dem GeR-Gesamtniveau der Lernenden vorliegt (p=0.187697, s.o.). Das Ergebnis des Modells ist in der nachfolgenden Abbildung (Abb. 47) visualisiert.

²⁹⁵ Aufgrund des Anspruchs einer konsistenten Visualisierung der hier vorgestellten Regressionsmodelle war eine Änderung der Reihenfolge der beiden Variablen notwendig. Damit weicht sie von der Reihenfolge entsprechender Variablen in RM2 ab. Dies hat jedoch keine Auswirkungen auf die Ergebnisse des Regressionsmodells.

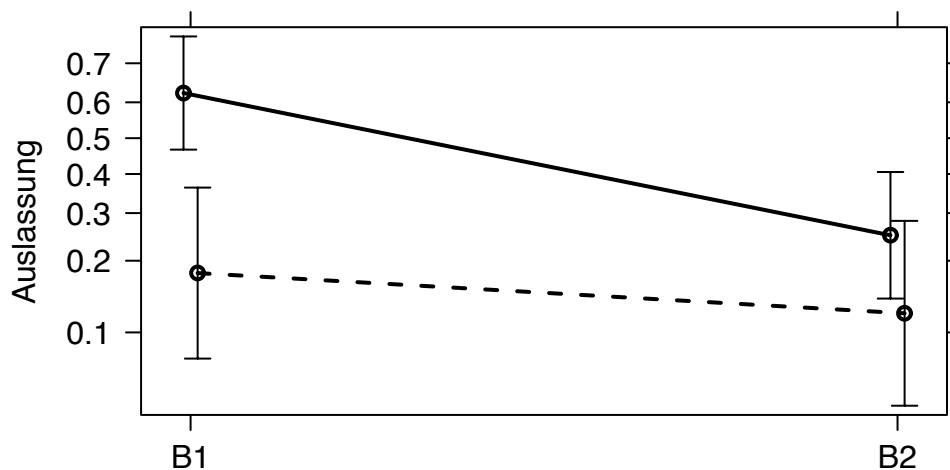


Abbildung 47: Fehlertyp Auslassung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau: Ergebnisse der Regressionsmodelle RM3, desemP = durchgezogene Linie, semP = gestrichelte Linie.

Die fehlende Interaktion zwischen den beiden Variablen wird in Abbildung 47 deutlich: Die Wahrscheinlichkeit für den Fehlertyp Auslassung ist für desemP (*an/auf*) in beiden untersuchten GeR-Gesamtniveaus B1 und B2 höher als für semP (*mit/für*), wobei der Unterschied im GeR-Gesamtniveau B1 deutlich größer ist als im GeR-Gesamtniveau B2 (s. entsprechenden Markierungen der Standardabweichungen der jeweiligen Mittelwerte in Abb. 47). Weiterhin zeigt die Abbildung eine Annäherung der Wahrscheinlichkeiten für den Fehlertyp Auslassung bei desemP und semP mit steigendem GeR-Gesamtniveau: Für desemP bedeutet dies eine starke Abnahme entsprechender Wahrscheinlichkeit (von ca. 60% auf ca. 25%). Auch bei semP lässt sich eine Abnahme der Wahrscheinlichkeit für den Fehlertyp Auslassung feststellen, allerdings handelt es sich um eine vergleichsweise geringe Abnahme (von ca. 20% auf ca. 13%). Dennoch bleibt hinsichtlich der Wahrscheinlichkeit für den Fehlertyp Auslassung ein deutlicher Unterschied zwischen semP und desemP bestehen.

Da im Regressionsmodell RM3 keine signifikante Interaktion zwischen dem semantischen Gehalt der Präposition und dem GeR-Gesamtniveau ermittelt werden konnte, wird das Regressionsmodell entsprechend modifiziert und lediglich die Einzeleffekte beider Variablen berücksichtigt. Das modifizierte Modell (RM3a) ist unten mit seinem Output angegeben und in der nachfolgenden Abbildung (Abb. 48) visualisiert.

RM3a:

Auslassung ~ Praeposition + Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa")

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.3589	0.3012	1.192	0.233415
PraepositionsemP	-1.5655	0.4424	-3.539	0.000402
Kompetenzniveau.opB2	-1.2728	0.4073	-3.125	0.001777

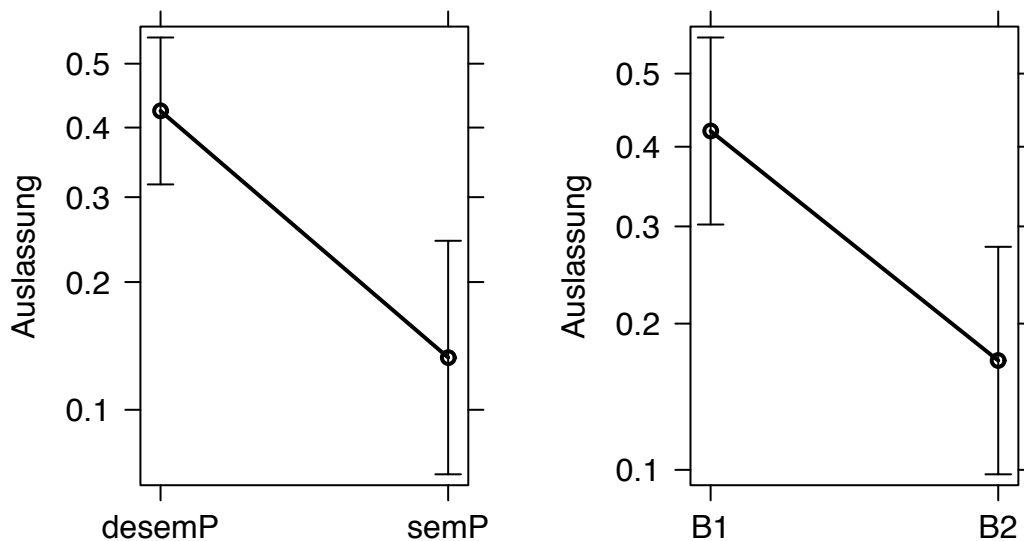


Abbildung 48: Fehlertyp Auslassung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM3a.

Die Ergebnisse des modifizierten Regressionsmodells (RM3a) (Abb. 48) zeigen, dass sowohl für den semantischen Gehalt der Präposition als auch für das GeR-Gesamtniveau signifikante Einzeleffekte auf den Fehlertyp Auslassung ermittelt werden können ($p=0.000402$ und $p=0.001777$, s. Modell-Output oben). Das bedeutet, dass beide Variablen unabhängig voneinander die Wahrscheinlichkeit des Fehlertyps Auslassung in Objekt-PP beeinflussen: Einerseits hat der semantische Gehalt der Präposition einen signifikanten Effekt und zwar dergestalt, dass desemP insgesamt eine signifikant höhere Wahrscheinlichkeit für Auslassungen von Präpositionen aufweisen als semP (ca. 43% zu ca. 15%, s. Abb. 48). Andererseits hat auch das GeR-Gesamtniveau der Lernenden einen signifikanten Effekt, und zwar dahingehend, dass die Wahrscheinlichkeit für den Fehlertyp Auslassung mit steigendem GeR-Gesamtniveau der Lernenden insgesamt signifikant abnimmt (ca. 43% auf B1 und ca. 18% auf B2, s. Abb. 48). Zusammenfassend lässt sich somit festhalten, dass Auslassungen mit steigendem Kompetenzniveau (von B1 zu B2) signifikant abnehmen, jedoch für desemP (im Vergleich zu semP) insgesamt signifikant wahrscheinlicher sind/bleiben. Die Evaluation des Modells (RM3a) zeigt einen R^2 -Wert von 0.1843438, d. h. das Modell kann die vorliegende Datenverteilung zu ca. 18,4% erklären.

Die signifikant höhere Wahrscheinlichkeit für präpositionale Auslassungen in Objekt-PP mit desemP kann vermutlich auf die Ausprägung ihres semantischen Gehalts zurückgeführt werden. Aktuellen Ansätzen zum L2-Erwerb (Ellis 2002: 175, Ellis 2006b, Mitchell/Myles/Marsden 2013: 105) folgend, werden Einheiten, die redundant für das Verständnis einer Äußerung sind, erst spät erworben. Da der semantische Gehalt von desemP in Objekt-PP reduziert ist, stehen diese mit einer geringe(re)n semantischen Funktion bzw. kommunikativen Relevanz in Verbindung. Dies erklärt, warum entsprechende Präpositionen tendenziell ausgelassen werden. Mit zunehmendem Kompetenzniveau wird es jedoch möglich, kommunikativ redundante Einheiten ausreichend zu verarbeiten

(Mitchell/Myles/Marsden 2013: 115) und sie damit in das sprachliche Wissen zu integrieren, was wiederum die abnehmende Wahrscheinlichkeit für Auslassungen von desemP erklärt.

Das Regressionsmodell (RM4) für den Fehlertyp Wahl ist parallel zum Regressionsmodell für den Fehlertyp Auslassung aufgebaut und ist im Folgenden mit seinem Output dargestellt. Im ersten Modell wird überprüft, ob eine Interaktion zwischen dem semantischen Gehalt der Präposition (Praeposition) und dem GeR-Gesamtniveau (Kompetenzniveau.op) der Lernenden vorliegt. Das Ergebnis des Modells ist in Abbildung 49 visualisiert.

RM4:

Wahl ~ Kompetenzniveau.op*Praeposition + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.2368	0.3786	-3.266	0.00109
Kompetenzniveau.opB2	1.4374	0.4943	2.908	0.00364
PraepositionsemP	-0.8835	0.7188	-1.229	0.21903
Kompetenzniveau.opB2:PraepositionsemP	-0.4566	0.8847	-0.516	0.60578

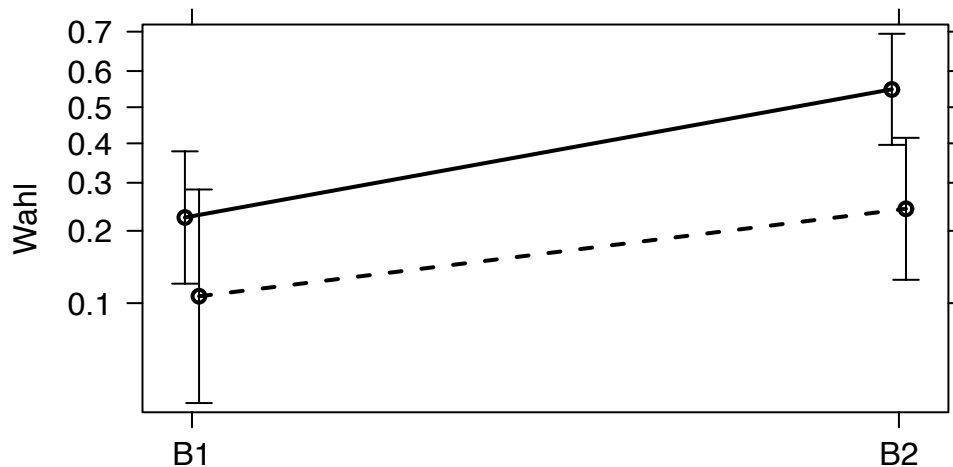


Abbildung 49: Fehlertyp Wahl im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau. Ergebnis des Regressionsmodells RM4, desemP = durchgezogene Linie, semP = gestrichelte Linie.

Wie aus den Werten des Modell-Outputs hervorgeht, liegt zwischen den beiden berücksichtigten Variablen keine Interaktion vor ($p=0.60578$, s.o.), d. h. die Variablen beeinflussen sich nicht gegenseitig.

Abbildung 49 zeigt eine konstant höhere Wahrscheinlichkeit für den Fehlertyp Wahl in Objekt-PP bei desemP (*an/auf*) im Vergleich zu semP (*mit/für*). Das bedeutet, dass die inkorrekte Wahl von Präpositionen eher bei desemP zu beobachten ist. Der Unterschied in der Wahrscheinlichkeit für den Fehlertyp Wahl ist dabei auf dem GeR-Gesamtniveau B2 größer als auf dem GeR-Gesamtniveau B1. Auf dem GeR-Gesamtniveau B2 lässt sich für desemP eine

deutlich höhere Wahrscheinlichkeit für wahlbedingte Fehler im Präpositionsgebrauch ermitteln. Weiterhin lässt sich feststellen, dass die Wahrscheinlichkeit für wahlbedingte Präpositionsfehler unabhängig vom semantischen Gehalt der Präposition zwischen den Kompetenzniveaus B1 und B2 ansteigt: Bei semP von ca. 10% auf ca. 25% und bei desemp von ca. 22% auf ca. 55%, d. h. die Wahrscheinlichkeit für wahlbedingte Präpositionsfehler nimmt mit steigendem Kompetenzniveau der DaF-Lernenden insgesamt zu.

Das Ergebnis des Regressionsmodells RM4 zeigt, wie oben erwähnt, keine Interaktion zwischen dem semantischen Gehalt der Präposition und dem GeR-Gesamtniveau. In einem modifizierten Regressionsmodell (RM4a) werden somit lediglich die Einzeleffekte des semantischen Gehalts der Präposition und des GeR-Gesamtniveaus berücksichtigt. Das modifizierte Modell (RM4a) ist im Folgenden mit seinem Output angegeben und in Abbildung 50 visualisiert.

RM4a:

Wahl ~ Praeposition + Kompetenzniveau.op + (1|Lerner), data = dat.subset, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.1578	0.3392	-3.413	0.000643
PraepositionsemP	-1.1926	0.4235	-2.816	0.004858
Kompetenzniveau.opB2	1.3017	0.4135	3.148	0.001642

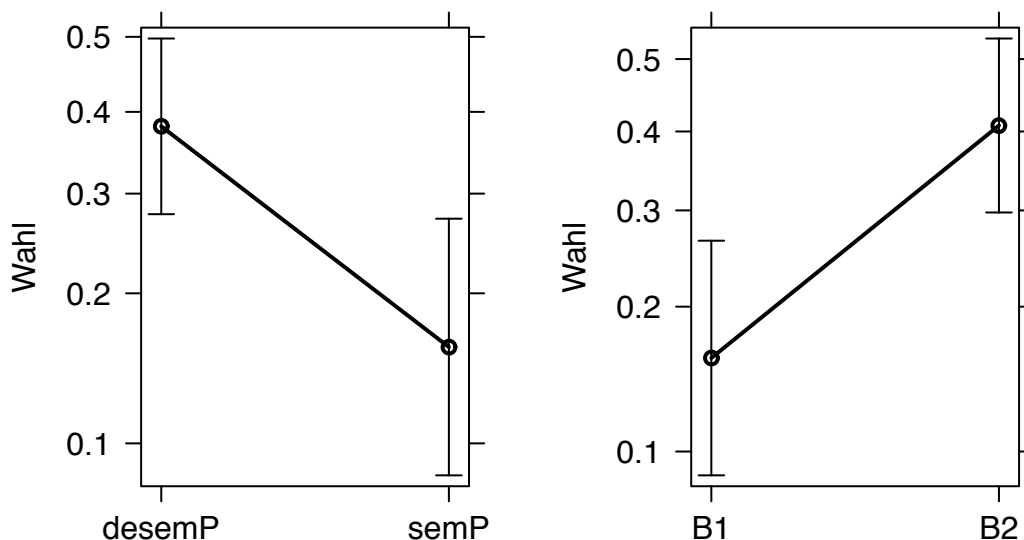


Abbildung 50: Fehlertyp Wahl im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemp) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM4a.

Ähnlich wie beim Fehlertyp Auslassung lässt sich auch beim Fehlertyp Wahl jeweils ein Einzeleffekt des semantischen Gehalts der Präposition als auch ein Einzeleffekt des GeR-Gesamtniveaus ermitteln ($p=0.004858$ und $p=0.001642$, s. Modell-Output oben). Diese Einzeleffekte sind in Abbildung 50 illustriert. Die Abbildung zeigt eine signifikant höhere Wahrscheinlichkeit für den Fehlertyp Wahl bei desemp im Vergleich zu semP (ca. 40% zu ca.

17%), und das unabhängig vom GeR-Gesamtniveau. Zusätzlich hat das GeR-Gesamtniveau einen Einfluss auf die Wahrscheinlichkeit für den Fehlertyp Wahl, und zwar unabhängig vom semantischen Gehalt der Präposition: Ein höheres GeR-Gesamtniveau bedingt den Daten nach einen signifikanten Anstieg in der Wahrscheinlichkeit für den Fehlertyp Wahl. Das bedeutet, dass wahlbedingte Präpositionsfehler mit steigendem GeR-Gesamtniveau zunehmen.

Wie in der Auswertung zur Korrektheit im Präpositionsgebrauch deutlich wurde, nimmt die Korrektheit im Präpositionsgebrauch mit steigendem GeR-Gesamtniveau insgesamt zu. Zugleich wurde in der Auswertung zum Fehlertyp Auslassung deutlich, dass mit steigendem GeR-Gesamtniveau Auslassungen von Präpositionen abnehmen. Die Zunahme wahlbedingter Fehler bei gleichzeitiger Abnahme des fehlerhaften Präpositionsgebrauchs (und Abnahme von präpositionalen Auslassungen) spiegelt insgesamt einen Lernfortschritt wider. DaF-Lernende zeigen mit steigendem GeR-Gesamtniveau zwar noch Unsicherheiten in der Wahl der entsprechenden Präposition, jedoch realisieren sie Präpositionen in zielsprachlich korrekten Kontexten.

Die Evaluation des Modells RM4a ergibt einen Wert von $R^2=0.1350049$, d. h. das Modell kann die Variation im vorliegenden Datensatz zu ca. 13,5% erklären. Damit liegt ein etwas niedrigerer Wert vor als bei der Evaluation des Modells zum Fehlertyp Hinzufügung bzw. Auslassung. Vergleicht man die R^2 -Werte der finalen Regressionsmodelle der hier untersuchten Fehlertypen Hinzufügung, Auslassung und Wahl miteinander, so zeigt das Regressionsmodell zum Fehlertyp Hinzufügung (RM2b) den höchsten R^2 -Wert ($R^2=0.2755648$), gefolgt vom R^2 -Wert des Regressionsmodells zum Fehlertyp Auslassung (RM3a) ($R^2=0.1843438$) und dem R^2 -Wert des Regressionsmodells zum Fehlertyp Wahl (RM4a) ($R^2=0.1350049$). Basierend auf diesen Werten kann geschlussfolgert werden, dass das Regressionsmodell zum Fehlertyp Hinzufügung (RM2b) dasjenige Modell ist, mit dem die vorliegende Varianz im Datensatz statistisch betrachtet am besten modelliert werden kann. Jedoch ist dabei zu beachten, dass für die Fehlertypen Wahl und Auslassung eine reduzierte Datengrundlage vorlag. Der vergleichsweise niedrigere R^2 -Wert könnte somit auch darin begründet sein.

Zusammenfassung

Auf Basis der statistischen Auswertungen zu den Fehlertypen im Präpositionsgebrauch kann die zweite Leitfrage folgendermaßen beantwortet werden:

2. Welche Fehlertypen treten bei *semantisierten Präpositionen* häufig auf und (wie) unterscheiden sich diese von den Fehlertypen im Gebrauch von *desemantisierten Präpositionen* in Objekt-PP?

Die Fehlertypen Hinzufügung, Auslassung und Wahl treten sowohl bei semantisierten Präpositionen (semP, hier: *an/auf*) als auch bei desemantisierten Präpositionen (desemP, hier: *mit/für*) in Objekt-PP auf. Das heißt, DaF-Lernende realisieren Präpositionen in Kontexten, in denen zielsprachlich keine Präpositionen bzw. PP gefordert werden, sie lassen zielsprachlich geforderte Präpositionen ganz aus oder sie wählen zielsprachlich inkorrekte Präpositionen. Die Auswertungen machen jedoch deutlich, dass abhängig vom semantischen Gehalt der Präposition unterschiedliche Fehlertypen erwartet werden können. So zeigen semP eine

signifikant höhere Wahrscheinlichkeit für den Fehlertyp Hinzufügung im Vergleich zu *desemP* während *desemP* eine höhere Wahrscheinlichkeit für die Auslassung von Präpositionen sowie für die inkorrekte Wahl von Präpositionen aufweisen. Insgesamt zeigen die statistischen Auswertungen, dass der semantische Gehalt der Präposition in Objekt-PP einen Einfluss auf den spezifischen Fehlertyp im Präpositionsgebrauch hat.

Die anfangs formulierten Hypothesen (3./4.) können den Ergebnissen nach teilweise bestätigt werden: Wie angenommen, zeigen *desemP* (*an/auf*) in Objekt-PP eine signifikant höhere Wahrscheinlichkeit für Auslassungen als *semP* (*mit/für*) (3. Hypothese). In Bezug auf den Fehlertyp Hinzufügung wurden keine Unterschiede erwartet (4. Hypothese), jedoch zeigen die quantitativen Auswertungen, dass hier deutliche Unterschiede bei *semP* und *desemP* ermittelt werden können: Für *semP* zeigt sich eine signifikante höhere Wahrscheinlichkeit für inkorrekte Hinzufügungen im Vergleich zu *desemP*.

Weiterhin kann für Auslassungen von Präpositionen in Objekt-PP festgestellt werden, dass die Wahrscheinlichkeit für diesen Fehlertyp mit steigendem GeR-Gesamtniveau signifikant abnimmt, was ebenfalls entsprechend vermutet wurde (5. Hypothese). Für Hinzufügungen kann solch ein Effekt nicht ermittelt werden: Hier zeigen sich unabhängig vom untersuchten GeR-Gesamtniveau ähnliche Wahrscheinlichkeiten für diesen Fehlertyp. Die statistischen Auswertungen verdeutlichen somit, dass die Wahrscheinlichkeit für den Fehlertyp Auslassung anhand des semantischen Gehalts der Präposition sowie anhand des GeR-Gesamtniveaus der Lernenden prognostiziert werden kann, beide Variablen haben folglich unabhängig voneinander einen Effekt. Dahingehen kann der Fehlertyp Hinzufügung allein auf Basis des semantischen Gehalts der Präposition prognostiziert werden, für das GeR-Gesamtniveau der Lernenden konnte hier kein Einfluss ermittelt werden. Für den Fehlertyp Wahl wurden ebenfalls keine spezifischen Annahmen formuliert und von ähnlichen Ausprägungen unabhängig vom semantischen Gehalt der Präposition ausgegangen (4. Hypothese). Diese Hypothese konnte nicht bestätigt werden. Die Auswertungen zeigen, dass für *desemP* eine signifikant höhere Wahrscheinlichkeit für diesen Fehlertyp ermittelt werden konnte als für *semP*. Wie anfangs angenommen (Hypothese 5), lässt sich jedoch zusätzlich feststellen, dass die Wahrscheinlichkeit für diesen Fehlertyp unabhängig vom semantischen Gehalt der Präposition mit steigendem GeR-Gesamtniveau zunimmt.

Zusammenfassend lässt sich somit festhalten, dass DaF-Lernende bei der Verwendung von *desemP* (*an/auf*) in Objekt-PP eine vergleichsweise höhere Wahrscheinlichkeit für Auslassungen von Präpositionen und wahlbedingte Präpositionsfehler zeigen, während sie bei *semP* (*mit/für*) eine signifikant höhere Wahrscheinlichkeit für inkorrekte Hinzufügungen von Präpositionen aufweisen. In allen Fällen hat der semantische Gehalt der Präposition in Objekt-PP einen Effekt auf den Fehlertyp im Präpositionsgebrauch.

10.6 Zusammenfassung und Diskussion der Gesamtergebnisse

In der vorliegenden Fallstudie wurde der Einfluss des semantischen Gehalts von Präpositionen auf den Gebrauch von Präpositionen durch DaF-Lernende untersucht. Der Fokus der Fallstudie lag auf Objekt-PP mit den semantisierten Präpositionen *mit/für* (*semP*) und den desemantisierten Präpositionen *an/auf* (*desemP*). Die Form-Funktionsverbindungen von *semP* und *desemP* in Objekt-PP unterscheiden sich hinsichtlich ihrer Zuverlässigkeit, die sich

im unterschiedlichen semantischen Gehalt entsprechender Präpositionen in Objekt-PP manifestiert: SemP weisen eine zuverlässige Form-Funktionsverbindung auf, da ihr semantischer Gehalt in Objekt-PP tendenziell ‚durchsichtig‘ und in- und außerhalb von Objekt-PP relativ konstant ist. DesemP dahingegen repräsentieren weniger zuverlässige Form-Funktionsverbindungen, da bei diesen keine semantische Durchsichtigkeit in Objekt-PP vorliegt und ihre Bedeutungszüge in Objekt-PP i.d.R. nicht auf eine Grundbedeutung zurückgeführt werden können, d.h. in- und außerhalb von Objekt-PP weisen sie einen anderen semantischen Gehalt auf (vgl. Breindl 1989: 39). Die semantische Durchsichtigkeit bei semP geht einher mit einer semantischen Relevanz entsprechender Präpositionen, die bei desemP nicht (in gleicher Weise) gegeben ist.

In der vorliegenden Fallstudie wurde die Annahme überprüft, ob die unterschiedlichen Form-Funktionsverbindungen bei semP und desemP in Objekt-PP und damit einhergehend der semantische Gehalt von Präpositionen einen Einfluss auf die Gebrauchsschwierigkeiten entsprechender Präpositionen durch DaF-Lernende haben. Die Schwierigkeiten im Präpositionsgebrauch wurden hier durch die Korrektheit und die Fehlertypen im Präpositionsgebrauch operationalisiert. Als Datengrundlage diente das Lernerkorpus MERLIN bzw. eine Teilmenge bereits erhobener und aufbereiteter Daten aus der Fallstudie 2 (s. Kap. 9), der ebenfalls das Lernerkorpus MERLIN als Datengrundlage diente. Entsprechend der Fragestellung wurden die erhobenen PP-Kontexte mit semP und desemP manuell aufbereitet, wobei u.a. die syntaktische Funktion der PP und die Fehlertypen im Präpositionsgebrauch (fehler)annotiert wurden. Im Rahmen einer computerunterstützten Fehleranalyse und einer kontrastiven Interlanguage-Analyse wurden die Korrektheit im Präpositionsgebrauch und die Fehlertypen im Präpositionsgebrauch ermittelt und semP und desemP kontrastiert. Die Daten wurden mittels Regressionsanalysen (gemischte Modelle) ausgewertet. Es wurde überprüft, ob bzw. inwiefern der semantische Gehalt der Präposition einen Einfluss auf die Korrektheit und die Fehlertypen im Präpositionsgebrauch hat. Für die Auswertungen wurde das GeR-Gesamtniveau der Lernenden berücksichtigt. Die Gesamtergebnisse der Auswertung zeigen, dass sowohl die Korrektheit im Präpositionsgebrauch als auch der Fehlertyp im Präpositionsgebrauch vom semantischen Gehalt der Präposition beeinflusst werden. Insgesamt zeigen sich bei semP weniger Gebrauchsschwierigkeiten als bei desemP.

In der vorliegenden Fallstudie wurden 814 Objekt-PP mit semP und desemP von insgesamt 489 DaF-Lernenden ausgewertet. Die Auswertung der Korrektheit im Präpositionsgebrauch zeigt insgesamt, dass die im Datensatz enthaltenen Präpositionen bzw. PP überwiegend korrekt realisiert werden. Dennoch wird ein Unterschied zwischen semP und desemP deutlich: Die statistische Auswertung bestätigt die 1. Hypothese, dass semP eine signifikant höhere Wahrscheinlichkeit für korrekten Präpositionsgebrauch in Objekt-PP zeigen als desemP. In der vorliegenden Fallstudie kann unabhängig vom GeR-Gesamtniveau der Lernenden ein Effekt des semantischen Gehalts auf die Korrektheit im Präpositionsgebrauch ermittelt werden. Zusätzlich kann ebenfalls ein Effekt des GeR-Gesamtniveau festgestellt werden: Mit zunehmenden Kompetenzniveau steigt die Wahrscheinlichkeit für korrekten Präpositionsgebrauch, und das unabhängig vom semantischen Gehalt der Präposition. Diese Ergebnisse bestätigen damit die 2. Hypothese. Das bedeutet, dass DaF-Lernende mit

steigender Sprachkompetenz insgesamt sicherer im Gebrauch von Präpositionen in Objekt-PP werden, jedoch zeigen sie dabei größere Unsicherheiten im Gebrauch von desemP (*an/auf*). Der Effekt des semantischen Gehalts auf die Korrektheit im Präpositionsgebrauch ist auf dem Niveau elementarer Sprachverwendung (A2) sowie im höheren Kompetenzniveau selbstständiger Sprachverwendung (B2) am größten. Das bedeutet, dass die Unsicherheiten im Gebrauch von desemP (im Vergleich zu semP) am größten sind zu Beginn des L2-Erwerbs sowie in der Phase, in der das sprachliche Wissen weiter ausgebaut und der Sprachgebrauch ‚kreativer‘ wird²⁹⁶. In Bezug auf die Korrektheit im Präpositionsgebrauch zeigt sich insgesamt ein positiver Effekt zuverlässiger Form-Funktionsverbindungen (bei semP) bzw. ein negativer Effekt weniger zuverlässiger Form-Funktionsverbindungen (bei desemP) auf die Korrektheit im Präpositionsgebrauch in Objekt-PP: SemP gehen mit einer höheren Wahrscheinlichkeit für Korrektheit im Präpositionsgebrauch einher als desemP. Eine entsprechende Tendenz können auch Hufeisen/Gibson (2002) in ihrer Untersuchung feststellen (vgl. Kap. 4.1). Dieses Ergebnis der vorliegenden Fallstudie kann weiterhin als Hinweis darauf gedeutet werden, dass in Objekt-PP semP schneller erworben und beherrscht werden als desemP. Ein ähnliches Ergebnis findet sich auch bei Carroll/Becker (1993: 146f.) in ihrer Untersuchung zum Erwerb von Raumreferenzen: Sie stellen für erwachsene Lernende fest, dass die semantische Transparenz lexikalischer Formen (z. B. Präpositionen) das Erwerbstempo dieser beeinflusst und lexikalische Formen zunächst in prototypischen Kontexten erworben und verwendet werden²⁹⁷, d. h. zuverlässige Form-Funktionsverbindungen werden zunächst erlernt. Wenn man davon ausgeht, dass die Korrektheit im Präpositionsgebrauch den Grad der Beherrschung von Präpositionen und damit Tendenzen im Erwerb von Präpositionen widerspiegelt (vgl. Turgay 2011), dann können die Ergebnisse der vorliegenden Fallstudie als Hinweis darauf gesehen werden, dass solch ein Effekt ebenfalls in den Ergebnissen der vorliegenden Fallstudie beobachtet werden kann.

Für die Auswertung der Fehlertypen im Präpositionsgebrauch wurden ausschließlich die als inkorrekt annotierten Objekt-PP-Kontexte von desemP und semP ausgewertet. Das GeR-Gesamtniveau der Lernenden wurde für die Analyse berücksichtigt, wobei GeR-Gesamtniveaus mit sehr niedrigen bzw. fehlenden Kontexten bestimmter Fehlertypen von der Auswertung ausgeschlossen wurden. Die Fehlertypen Auslassung, Wahl und Hinzufügung können sowohl bei semP als auch bei desemP ermittelt werden. Es zeigen sich jedoch in Abhängigkeit des semantischen Gehalts der Präposition deutliche Unterschiede in der Wahrscheinlichkeit für die einzelnen Fehlertypen, und das nicht nur für den Fehlertyp Auslassung, wie eingangs angenommen wurde (3./4. Hypothese). Die Auswertung der Fehlertypen unterstreicht den Effekt unterschiedlicher Form-Funktionsverbindungen von semP und desemP in Objekt-PP: Für desemP (*an/auf*) kann unabhängig vom GeR-

²⁹⁶ Im GeR (Europarat 2001: 35) wird das Referenzniveau B2 u.a. dadurch definiert, dass der Lernende sich „spontan und fließend verständigen“ und sich zu einem breiten Themenspektrum äußern kann. Ich verstehe hier darunter, dass der Lernende ein sprachliches Wissen kreativ und eigenständig einsetzt und sprachlich in der Lage ist, bereits erworbenes Wissen zu transferieren.

²⁹⁷ Kidd/Cameron-Faulkner (2008) liefern ähnliche Beobachtungen zum muttersprachlichen Erwerb verschiedener Bedeutungen der englischen Präposition *with*.

Gesamtniveau der Lernenden eine signifikant höhere Wahrscheinlichkeit für die Fehlertypen Auslassungen und Wahl im Präpositionsgebrauch ermittelt werden als für semP (*mit/für*). Somit kann ein Effekt des semantischen Gehalts der Präposition auf den Fehlertyp im Präpositionsgebrauch festgestellt werden. Dieses Ergebnis könnte ebenfalls auf die Zuverlässigkeit der jeweiligen Form-Funktionsverbindungen zurückzuführen sein: Da semP (*mit/für*) in Objekt-PP „relativ konstante Bedeutungszüge“ (Breindl 1989: 39) aufweisen und damit eine zuverlässige Form-Funktionsverbindung repräsentieren, kommt es hier zu deutlich weniger Auslassungen von Präpositionen oder wahlbedingten Fehlern im Präpositionsgebrauch. Im Gegensatz dazu ist die Bedeutung von desemP (*an/auf*) in Objekt-PP nicht gut isolierbar bzw. auf die Grundbedeutung dieser Präpositionen außerhalb von Objekt-PP zurückführbar. Hier zeigen sich somit weniger zuverlässige Form-Funktionsverbindungen und dies kann die vergleichsweise signifikant höhere Wahrscheinlichkeit für Auslassungen sowie wahlbedingte Fehler im Präpositionsgebrauch bei desemP erklären (vgl. auch Viorel 1973 für eine ähnliche Beobachtung zu wahlbedingten Fehlern im Präpositionsgebrauch, Kap. 4.1). Parallel zum Effekt des semantischen Gehalts der Präposition auf die Fehlertypen Auslassung und Wahl zeigt sich ebenfalls ein Effekt des GeR-Gesamtniveaus: Unabhängig vom semantischen Gehalt der Präposition nehmen Auslassungen von Präpositionen in Objekt-PP mit steigendem GeR-Gesamtniveau ab, wahlbedingte Präpositionsfehler nehmen dahingegen zu. Damit bestätigen die Ergebnisse die 5. Hypothese. In der Zunahme wahlbedingter Präpositionsfehler bei gleichzeitiger Abnahme von Präpositionsauslassungen manifestiert sich insgesamt ein Lernfortschritt im Gebrauch von Präpositionen, und das unabhängig vom semantischen Gehalt der Präposition.

Beim Fehlertyp Hinzufügung zeigt sich ebenfalls ein deutlicher Effekt des semantischen Gehalts der Präposition unabhängig vom GeR-Gesamtniveau der Lernenden, denn semP (*mit/für*) zeigen eine signifikant höhere Wahrscheinlichkeit für diesen Fehlertyp als desemP (*an/auf*). Das bedeutet, dass DaF-Lernende dazu tendieren, die Realisierung von semP (*mit/für*) zu übergeneralisieren und semP in Kontexten zu realisieren, in denen zielsprachlich keine Präpositionen bzw. PP gefordert sind. Einige der Kontexte, in denen Lernende semP hinzufügen, illustrieren, dass die zuverlässigen Form-Funktionsverbindungen von semP auf semantisch ähnliche Kontexte übertragen werden, auch wenn zielsprachlich an entsprechenden Stellen keine PP zur Realisierung der Funktion notwendig ist, wie z. B. die Realisierung von *für*-PP anstelle eines Dativobjekts zur Anzeige eines/einer Nutznießenden bzw. eines/einer Rezipienten/-in (s. auch Pfaff 1984 und Kaltenbacher/Klages 2006, s. Kap. 4.1). Ein Effekt des semantischen Gehalts auf den Fehlertyp Hinzufügung kann unabhängig vom GeR-Gesamtniveau der Lernenden ermittelt werden; ein zusätzlicher Effekt des GeR-Gesamtniveau liegt hier, anders als bei den anderen Fehlertypen, nicht vor. Das bedeutet, dass die Wahrscheinlichkeit des Fehlertyps Hinzufügung lediglich auf Basis des semantischen Gehalts der Präposition gut vorhergesagt werden kann. Weitere Untersuchungen wären hier notwendig, um den Einfluss des Semantischen Gehalts der Präposition genauer herauszuarbeiten. Vor allem bei PP mit *für* wäre es denkbar, dass diese aufgrund der semantischen Nähe zur Dativ-NP eine Sonderstellung in der Gruppe der semantisierten Präpositionen einnehmen.

Die Ergebnisse der Fallstudie liefern erste, empirisch fundierte Erkenntnisse zum Einfluss des semantischen Gehalts von Präpositionen in Objekt-PP auf den Gebrauch von Präpositionen durch DaF-Lernende. Die Auswertungen zeigen unterschiedliche Gebrauchs(un)sicherheiten und spezifische Fehlertypen im Präpositionsgebrauch in Objekt-PP mit desemP (*an/auf*) und semP (*mit/für*), d. h. in Abhängigkeit vom semantischen Gehalt der Präposition. Geht man davon aus, dass die in Kapitel 1.2 vorgestellte Kasushierarchie, nach der Objekt-PP die rangniedrigste Stellung einnehmen, sich ebenfalls in der Erwerbsreihenfolge entsprechender Formen manifestiert (vgl. Wegener 1995a: 352), so müssten Objekt-PP mit semP hierarchisch höher anzusetzen sein als Objekt-PP mit desemP. Die Ergebnisse zum Einfluss des semantischen Gehalts der Präposition in Objekt-PP auf die Korrektheit im Präpositionsgebrauch machen deutlich, dass Objekt-PP mit desemP Deutschlernende vor andere Herausforderungen stellen als Objekt-PP mit semP: DaF-Lernende zeigen hier quantitativ (Korrektheit) und qualitativ (Fehlertypen) betrachtet unterschiedliche Gebrauchsschwierigkeiten.

11. Zusammenfassung Teil II und Ausblick

Ziel der vorliegenden Arbeit war es, im Lernerkorpus MERLIN erwerbstheoretisch informierte, quantitative Fallstudien zu ausgewählten Einflussvariablen im Gebrauch von Präpositionen und PP durch DaF-Lernende durchzuführen und damit u.a. das Potenzial von Lernerkorpora für die quantitative (DaF-)Spracherwerbsforschung zu illustrieren. Die Dissertation ist in der Lernerkorpusforschung an der Schnittstelle zur Z/FSE-Forschung zu verorten und orientiert sich an aktuellen gebrauchsbasierten Ansätzen zum Z/FSE. Die sich aus diesen Ansätzen zum Zweit- bzw. Fremdspracherwerb (L2-Erwerb) ergebenden Hypothesen wurden anhand quantitativer Auswertungen in einem Lernerkorpus überprüft. Im Fokus der empirischen Untersuchungen stand der Einfluss von Form-Funktionsverbindungen sprachlicher Einheiten, die im Kontext aktueller Spracherwerbstheorien einen zentralen Stellenwert einnehmen. Hier ist vor allem der Einfluss bereits bestehender Form-Funktionsverbindungen (z. B. aus der Erstsprache) (FS 1) zu nennen sowie der Einfluss der Zuverlässigkeit von Form-Funktionsverbindungen in der Zielsprache Deutsch (FS 2 und FS 3). Im Einzelnen wurden in der Dissertation die folgenden ausgewählten Einflussvariablen untersucht:

- Erstsprache (L1) der Lernenden
- Syntaktische Funktion der PP
- Semantischer Gehalt der Präposition

Jede Einflussvariable wurde in einer empirischen Fallstudie im Lernerkorpus MERLIN untersucht, wobei zusätzlich die Einflussvariable Sprachkompetenz in jeder Fallstudie berücksichtigt wurde.

Die einzelnen Fallstudien nehmen die in Kapitel 1 ausgearbeiteten Charakteristiken von deutschen Präpositionen und PP aus der Perspektive aktueller Spracherwerbstheorien in den Fokus (s. Kap. 2) und berücksichtigen die in Kapitel 4 herausgearbeiteten methodologischen und inhaltlichen Forschungslücken zum Erwerb und Gebrauch von

Präpositionen und PP durch Deutschlernende. Die übergeordnete Fragestellung der Dissertation bezog sich vor allem auf die Gebrauchsschwierigen, die sich in diesem Bereich zeigen. Für die empirischen Untersuchungen wurden *Gebrauchsschwierigkeiten* operationalisiert durch die Korrektheit im Präpositionsgebrauch sowie durch die einzelnen Fehlertypen im Gebrauch von Präpositionen und PP. Die theoretischen Grundlagen der Dissertation wurden in Kapitel 5 ausführlich zusammengefasst.

Datengrundlage der einzelnen Fallstudien war das digitale, frei verfügbare Lernerkorpus MERLIN, das in Kapitel 6 ausführlich vorgestellt wurde. Der Schwerpunkt der Fallstudien im Lernerkorpus lag auf quantitativen Auswertungen unter Rückgriff auf zentrale Analysemethoden der Lernerkorpusforschung (computerunterstützte Fehleranalyse und kontrastive Interlanguage-Analyse) und statistischen Modellierungen zum Einfluss der ausgewählten Variablen auf den Gebrauch von Präpositionen und PP mittels logistischer Regressionsanalysen. Die Analysemethoden und statistischen Auswertungen wurden in Kapitel 7 im Detail erläutert.

Die zentralen Ergebnisse der Fallstudien zu ausgewählten Einflussvariablen und ihren Effekten auf die Korrektheit und die Fehlertypen im Gebrauch von Präpositionen und PP durch Deutschlernende wurden in den entsprechenden Kapiteln präsentiert und diskutiert (Kap. 8-10). In der nachfolgenden Tabelle (Tab. 70) sind die Ergebnisse kompakt zusammengefasst²⁹⁸.

Mithilfe der Ergebnisse können die eingangs formulierten übergeordneten Fragestellungen nun entsprechend beantwortet werden. Die übergeordneten Fragestellungen seien an dieser Stelle nochmals erwähnt:

Welche Variablen beeinflussen den Gebrauch von Präpositionen und Präpositionalphrasen durch Lernende des Deutschen als Fremdsprache?

Welche der ausgewählten Einflussvariablen bedingen (welche) Gebrauchsschwierigkeiten?

²⁹⁸ Die Variable *Sprachkompetenz* ist durch das GeR-Gesamtniveau repräsentiert. Pfeile in der Tabelle zeigen an, dass an entsprechender Stelle eine signifikant höhere Wahrscheinlichkeit ermittelt werden konnte. Beispiel: Die Ergebnisse von Fallstudie 1 zeigen, dass Lernende mit L1TA eine signifikant höhere Wahrscheinlichkeit für wahlbedingte Präpositionsfehler zeigen (im Vergleich zu Lernenden mit L1TV).

Fallstudie (FS)	Fallstudie 1 (FS1)	Fallstudie 2 (FS 2)	Fallstudie 3 (FS 3)
Einflussvariable	Erstsprache (L1) der Lernenden	Syntaktische Funktion der PP (SynF)	Semantischer Gehalt der Präposition (SemG)
Eingrenzung der Einflussvariable	L1TV (Ungarisch/Türkisch) vs. L1TA (Englisch/Italienisch)	Objekt-PP vs. adverbiale PP Präpositionen: <i>an/auf</i>	Semantisierte (semP) (<i>an/auf</i>) vs. desemantisierte (desemp) Präpositionen (<i>mit/für</i>) in Objekt-PP
Einfluss auf Korrektheit im Präpositionsgebrauch	<ul style="list-style-type: none"> Kein signifikanter Effekt der L1 Signifikanter Effekt des GeR-Gesamtniveaus 	<ul style="list-style-type: none"> Interaktion zwischen GeR-Gesamtniveau und SynF Signifikanter Effekt von SynF nur auf GeR-Gesamtniveau A2 	<ul style="list-style-type: none"> Signifikanter Effekt des SemG Signifikanter Effekt des GeR-Gesamtniveaus
Spezifizierung des Einflusses auf Korrektheit	<ul style="list-style-type: none"> Zunahme Korrektheit mit steigendem GeR-Gesamtniveau 	<ul style="list-style-type: none"> ↑ A2: Korrektheit adverbialer PP 	<ul style="list-style-type: none"> ↑ Korrektheit semP Zunahme Korrektheit mit steigendem GeR-Gesamtniveau
Einfluss auf Fehlertyp im Präpositionsgebrauch	<ul style="list-style-type: none"> Signifikanter Effekt der L1 Signifikanter Effekt des GeR-Gesamtniveaus 	<ul style="list-style-type: none"> Signifikanter Effekt der SynF Signifikanter Effekt des GeR-Gesamtniveaus 	<ul style="list-style-type: none"> Signifikanter Effekt des SemG Signifikanter Effekt des GeR-Gesamtniveaus
Spezifizierung des Einflusses auf den Fehlertyp im Präpositionsgebrauch	<ul style="list-style-type: none"> L1: <ul style="list-style-type: none"> ↑ L1TA: Fehler in Wahl der Präposition ↑ L1TV: Fehler in der Realisierung (Auslassung und Hinzufügung) von Präpositionen GeR-Gesamtniveau: Mit Zunahme des GeR-Gesamtniveaus: <ul style="list-style-type: none"> Abnahme Realisierung-Fehler (bedingte) Zunahme Wahl-Fehler 	<ul style="list-style-type: none"> SynF: <ul style="list-style-type: none"> ↑ Adv. PP: Fehler in Wahl der Präposition ↑ Objekt-PP: Fehler in Auslassung der Präposition GeR-Gesamtniveau: Mit Zunahme des GeR-Gesamtniveaus: <ul style="list-style-type: none"> Abnahme Auslassung-Fehler Zunahme Wahl-Fehler 	<ul style="list-style-type: none"> SemG: <ul style="list-style-type: none"> ↑ desemp: Fehler in Auslassung und Wahl der Präposition ↑ semP: Fehler in Hinzufügung von Präpositionen GeR-Gesamtniveau: Mit Zunahme des GeR-Gesamtniveaus: <ul style="list-style-type: none"> Abnahme Auslassung-Fehler Zunahme Wahl-Fehler Kein Effekt auf Fehler in Hinzufügung

Tabelle 70: Fallstudien im Lernerkorpus MERLIN: Zentrale Ergebnisse zur Korrektheit und Fehlertypen im Präpositionsgebrauch.

Die drei Fallstudien (vgl. Tab. 70) fokussierten jeweils eine ausgewählte Einflussvariable im Gebrauch von Präpositionen und PP, für die ich in der bisherigen Forschung entsprechende methodologische und inhaltliche Forschungslücken herausgearbeitet habe (s. Kap. 5). Die Einflussvariablen umfassten interlinguale sowie intralinguale Variablen im L2-Erwerb. Als interlinguale Einflussvariable stand die typologische Nähe der L1 zur Zielsprache Deutsch im Fokus (FS 1, Kap. 8). Als intralinguale Einflussvariablen wurden die syntaktische Funktion der PP (FS 2, Kap. 9) sowie der semantische Gehalt der Präposition (FS 3, Kap. 10) untersucht. Zusätzlich berücksichtigt wurde in jeder Fallstudie die Sprachkompetenz bzw. das Sprachkompetenzniveau der DaF-Lernenden.

Die Ergebnisse der einzelnen Fallstudien (Tab. 70) zeigen, dass sich die Schwierigkeiten im Gebrauch von Präpositionen und PP durch DaF-Lernende nicht primär in der Korrektheit im Präpositionsgebrauch widerspiegeln – de facto zeigen die DaF-Lernenden in den untersuchten Daten insgesamt eine recht hohe Korrektheit im Präpositionsgebrauch. Die Gebrauchsschwierigkeiten manifestieren sich vorwiegend in unterschiedlichen Fehlertypen im Präpositionsgebrauch. Die Gebrauchsschwierigkeiten sind demnach nicht quantitativer (Korrektheit), sondern eher qualitativer Natur (Fehlertypen). Im Einzelnen konnten in den Fallstudien die Relevanz folgender Einflussvariablen im Gebrauch von Präpositionen und PP nachgewiesen werden (vgl. auch Tab. 70): Die L1 der Lernenden, die syntaktische Funktion der PP (Objekt-PP vs. adverbiale PP) sowie der semantische Gehalt der Präposition (semP vs. desemp in Objekt-PP). Hier zeigt sich jeweils ein signifikanter Effekt auf die Fehlertypen im nicht-zielsprachlichen Gebrauch von Präpositionen und PP durch DaF-Lernende. In Bezug auf die Sprachkompetenz zeigt sich zusätzlich, dass die Schwierigkeiten im Gebrauch von Präpositionen und PP in den meisten Fällen unabhängig von der jeweils anderen untersuchten Einflussvariable mit steigender Sprachkompetenz der Lernenden abnehmen und somit die Sicherheit im Gebrauch von Präpositionen und PP zunimmt.

Aus der Perspektive aktueller spracherwerbstheoretischer Annahmen gebrauchsbasierter Ansätze liefern die Fallstudien wertvolle Erkenntnisse zum Einfluss von Form-Funktionsverbindungen: Aus den Ergebnissen der einzelnen Fallstudien lässt sich jeweils ein signifikanter Einfluss der Zuverlässigkeit von Form-Funktionsverbindungen sprachlicher Strukturen in der Zielsprache Deutsch (FS 2 und FS 3) sowie der Einfluss bereits aufgebauter, verfestigter Form-Funktionsverbindungen (aus der L1) (FS 1) auf den Gebrauch zielsprachlicher Strukturen ableiten. Dabei wird deutlich, dass weniger zuverlässige Form-Funktionsverbindungen der Zielsprache Deutsch (= Objekt-PP in FS 2 und desemp in FS 3) nicht per se zu mehr Schwierigkeiten im Gebrauch von Präpositionen und PP führen, sondern andere Gebrauchsschwierigkeiten bedingen als zuverlässige Form-Funktionsverbindungen (= adverbiale PP in FS 2 und semP in FS 3). Dabei gehen weniger zuverlässige Form-Funktionsverbindungen mit häufigen Auslassungen von Präpositionen einher. Für die anderen untersuchten Fehlertypen (Wahl und Hinzufügung) lässt sich auf Basis der statistischen Auswertungen dahingegen kein eindeutiges, fallstudienübergreifendes Bild zeichnen. Bereits aufgebaute Form-Funktionsverbindungen aus der L1 zeigen den Auswertungen nach einen deutlichen Einfluss auf den Gebrauch von Präpositionen und PP in der Zielsprache Deutsch: So zeigen z.B. Lernende, die in ihrer L1 über Präpositionen und PP verfügen (und damit eine

Nähe zu entsprechenden Form-Funktionsverbindungen in der Zielsprache Deutsch aufweisen), eine signifikant niedrige Wahrscheinlichkeit für präpositionale Auslassungen und Hinzufügungen. Bei diesen Lernenden lässt sich jedoch eine signifikant höhere Wahrscheinlichkeit für wahlbedingte Fehler im Gebrauch von Präpositionen feststellen (vgl. Tab. 70).

Die Ergebnisse der Fallstudien spiegeln ebenfalls die wachsende Sicherheit von DaF-Lernenden im Gebrauch von Präpositionen und PP mit unterschiedlichen Form-Funktionsverbindungen wider: In (nahezu²⁹⁹) jeder Fallstudie zeigt sich ein signifikanter Einzeleffekt der Sprachkompetenz (GeR-Gesamtniveau) auf die Korrektheit im Präpositionsgebrauch, und das unabhängig von der jeweils untersuchten Einflussvariable. Das bedeutet, dass unabhängig von den hier untersuchten Einflussvariablen sich DaF-Lernende mit Zunahme der Sprachkompetenz dem zielsprachlichen Gebrauch von Präpositionen und PP stark annähern. Zudem kann in jeder Fallstudie unabhängig von der untersuchten Einflussvariable ein signifikanter Effekt der Sprachkompetenz auf die zentralen Fehlertypen im Gebrauch von Präpositionen und PP ermittelt werden: Im nicht-zielsprachlichen Gebrauch von Präpositionen und PP nehmen Auslassungen von Präpositionen mit Zunahme der Sprachkompetenz insgesamt ab, während Fehler in der Wahl der zielsprachlich korrekten Präposition mit steigender Sprachkompetenz zunehmen. Damit spiegeln die Fehlertypen die zunehmende Sicherheit im Präpositionsgebrauch wider: Präpositionen werden in zielsprachlich korrekten Kontexten zunehmend häufiger realisiert, jedoch bereitet die Wahl der korrekten Präposition scheinbar (noch) ein Problem. Die Ergebnisse zur Korrektheit im Gebrauch von Präpositionen und PP verdeutlichen jedoch, dass die Fehler im Präpositionsgebrauch insgesamt abnehmen.

Die einzelnen empirischen Fallstudien der Dissertation fokussieren einen Teilbereich des Gebrauchs von Präpositionen und PP durch Deutschlernende. Die Ergebnisse zu Einflussvariablen im Gebrauch von Präpositionen und PP durch DaF-Lernende sind daher auf diesen Teilbereich beschränkt. Die Einschränkung betrifft allen voran die Fokussierung auf den Gebrauch der Präposition (unabhängig von der Kasusreaktion) sowie die Anzahl an Einflussvariablen, die im Rahmen der Arbeit quantitativ untersucht wurden. Weiterhin waren ebenfalls Einschränkungen in den einzelnen Fallstudien notwendig. Diese Einschränkungen betreffen a) die ausgewählten Präpositionen (vor allem *an*, *auf*, *für*, *mit*), b) die ausgewählten syntaktischen Funktionen von PP (vor allem Objekt-PP und adverbiale PP) sowie c) die ausgewählten Erstsprachen der Lernenden (vor allem die typologisch verschiedenen L1 Türkisch und Ungarisch sowie die typologisch ähnlichen L1 Englisch und Italienisch). Ob und inwiefern die hier erlangten Erkenntnisse zu Einflussvariablen

- auf andere Präpositionen
- auf andere syntaktische Funktionen von PP
- auf andere typologisch verschiedene L1 bzw. typologisch ähnliche L1

übertragbar sind, gilt es, in weiteren, anknüpfenden Untersuchungen in Lernerkorpora empirisch zu überprüfen. Weiterhin können die im Rahmen dieser Dissertation erlangten

²⁹⁹ In Fallstudie 2 liegt eine Interaktion vor zwischen den Variablen *SynF* und *GeR-Gesamtniveau*.

Erkenntnisse lediglich in Bezug auf die analysierten Korpusdaten interpretiert werden und sind daher nicht zu verallgemeinern.

Die zentralen Einflussvariablen (Erstsprache der Lernenden, syntaktische Funktion der PP und semantischer Gehalt der Präposition) wurden in der vorliegenden Dissertation in separaten Fallstudien behandelt und daher getrennt voneinander untersucht. Die Frage, ob bzw. inwiefern sich die einzelnen Variablen in ihrem Effekt auf den Gebrauch von Präpositionen und PP durch DaF-Lernende gegenseitig beeinflussen, gilt es, in weiterführenden empirischen Untersuchungen in Lernerkorpora zu erörtern. Die vorliegende Dissertation kann hierbei als wegweisend betrachtet werden. Sie zeigt exemplarisch, wie Annahmen aus der Spracherwerbsforschung anhand von Daten aus einem digitalen Lernerkorpus überprüft und mögliche Einflussvariablen im Sprachgebrauch von Lernenden ermittelt werden können. Zusammenfassend betrachtet offenbaren die Ergebnisse der erwerbstheoretisch informierten, quantitativen Untersuchungen im Lernerkorpus MERLIN, dass die Einflussvariablen im Gebrauch von Präpositionen und PP durch DaF-Lernende sowohl auf die Merkmale der Zielsprache Deutsch zurückgeführt werden können als auch auf Merkmale der L1 der Lernenden. Damit bestätigen die Ergebnisse den Einfluss intralingualer sowie interlingualer Merkmale im L2-Erwerb. Die Untersuchungen bieten wertvolle empirisch fundierte und statistisch gestützte Erkenntnisse für die Lernerkorpusforschung an der Schnittstelle zur Zweit- und Fremdspracherwerbsforschung und illustrieren damit das Potenzial von Lernerkorpora als Ressource für die quantitative DaF-Spracherwerbsforschung.

Weiterhin bietet die Arbeit mit ihren Erkenntnissen zu spezifischen Fehlertypen im Präpositionsgebrauch in Abhängigkeit von bestimmten Variablen ebenfalls wertvolle Anknüpfungspunkte für die Schnittstelle zum Zweit- und Fremdsprachenunterricht sowie zur maschinellen Sprachverarbeitung: So können die Erkenntnisse der Fallstudien genutzt werden für die Konzipierung von Lehrmaterialien und/oder die Schwerpunktsetzung in der Vermittlung von Präpositionen und PP im DaZ-/DaF-Kontext. Rundell/Granger (2007) und Granger (2015: 491f.) zeigen z. B. für Englisch als Zielsprache, wie Erkenntnisse aus lernersprachlichen Analysen auf Basis von Lernerkorpusdaten die Inhalte von Nachschlagewerken und Lernerwörterbüchern optimieren können. Für den DaZ-/DaF-Bereich birgt ein entsprechender Ansatz ebenfalls großes Potenzial; aus den Ergebnissen dieser Dissertation kann z. B. abgeleitet werden, dass DaF-Lernende in Abhängigkeit ihrer L1 unterschiedliche lernunterstützende Hinweise zum Gebrauch von Präpositionen und PP benötigen bzw. unterschiedliche didaktische Schwerpunktsetzungen notwendig sind: Für Lernende mit Türkisch bzw. Ungarisch als L1 sind dies vor allem typische Gebrauchskontexte von Präpositionen und PP im Deutschen, für Lernende mit Englisch bzw. Italienisch dahingegen vor allem Präpositionsbedeutungen (und ggf. ihre anderssprachigen Äquivalente).

An der Schnittstelle zur maschinellen Sprachverarbeitung bietet die Dissertation hilfreiche Orientierungspunkte für die Entwicklung von Systemen zur automatischen Erkennung und Korrektur prototypischer Abweichungen im Präpositionsgebrauch bei DaF-Lernenden. Bisherige Ansätze versuchen auf Basis muttersprachlicher, redigierter Daten (s. z. B. Chodorow/Tetreault/Han 2007) oder Parallelkorpora (Graën/Schneider 2017) Abweichungen im lernersprachlichen Präpositionsgebrauch zu prognostizieren und diese

Informationen für die automatische Fehlererkennung bzw. -korrektur zu nutzen. Die auf diese Weise entwickelten Modelle beziehen keine lernersprachlichen Analysen mit ein, sind folglich nicht auf authentische Lernerproduktionen abgestimmt. Dies wirkt sich negativ auf die Leistung entsprechender Modelle aus (vgl. Chodorow/Tetreault/Han 2007: 28). In den einzelnen Fallstudien dieser Dissertation wurden authentische Lernerproduktionen ausgewertet. Die Ergebnisse liefern wertvolle Informationen zu relevanten, prototypischen Abweichungen im Gebrauch ausgewählter Präpositionen in der Zielsprache Deutsch (s. Tab. 70), die gewinnbringend in der Entwicklung entsprechender Systemen zur automatischen Fehlererkennung und -korrektur berücksichtigt werden können. Wie die Auswertungen der Fallstudien zeigen, betreffen die meisten Abweichungen im Gebrauch deutscher Präpositionen die Auslassung bzw. die Wahl von Präpositionen – ein automatisches Fehlererkennungssystem müsste somit in der Lage sein, vor allem diese Art der Abweichungen zuverlässig erkennen zu können.

Das Lernerkorpus MERLIN, das hier als Datengrundlage diente, ist eine wertvolle Ressource für Untersuchungen an der Schnittstelle zur DaF-Spracherwerbsforschung. Wie die empirischen Fallstudien zeigen, bietet MERLIN eine gute Datengrundlage für quantitative Auswertungen und Modellierungen möglicher Einflussvariablen im Gebrauch von Präpositionen und PP, auch wenn die vollständige Fehlerannotation aller Texte sowie eine homogenere Verteilung der Texte auf die einzelnen GeR-Gesamtniveaus ein Desiderat darstellt. Die deutsche Lernerkorpuslandschaft befindet sich derzeit³⁰⁰ noch im Ausbau; aktuelle Projekte (z. B. Wisniewski 2018a, Hirschmann/Nolda 2019) lassen jedoch auf eine blühende Zukunft deutscher Lernerkorpora hoffen, sodass in den nächsten Jahren davon auszugehen ist, dass einerseits die Anzahl deutscher Lernerkorpora ansteigen wird und andererseits korpuslinguistische Untersuchungen in deutschen Lernerkorpora zunehmen werden. Die vorliegende Dissertation liefert erste quantitative, erwerbstheoretisch informierte und korpusgestützte Erkenntnisse zu Einflussvariablen im Präpositionsgebrauch durch DaF-Lernende, für die bisher systematische Untersuchungen fehlten. Sie dient damit auch als Anstoß für weitere Arbeiten zur Erforschung von Einflussvariablen im Spracherwerb und Sprachgebrauch auf der Basis von Lernerkorpora.

³⁰⁰ Stand: April 2019.

Abkürzungsverzeichnis

AL	Anzahl unterschiedlicher Lernertexte
DaF	Deutsch als Fremdsprache
DaZ	Deutsch als Zweitsprache
desemP	desemantisierte Präposition(en)
E-Valbu	Das Elektronische Valenzwörterbuch deutscher Verben
GeR	Gemeinsamer europäischer Referenzrahmen für Sprachen
L1	Erstsprache
L1TA	Dem Deutschen typologisch ähnliche Erstsprache
L1TV	Vom Deutschen typologisch verschiedene Erstsprache
L2	Zweit- o. Fremdsprache bzw. Zielsprache
NLP	Natural language processing (dt. maschinelle Sprachverarbeitung)
PP	Präpositionalphrase(n)
RM	Regressionsmodell
semP	semantisierte Präposition(en)
TA	Dem Deutschen typologisch ähnlich
TV	Vom Deutschen typologisch verschieden
Z/FSE	Zweit-/Fremdspracherwerb
ZH	Zielhypothese

Abbildungsverzeichnis

Abbildung 1: Kasushierarchie im Deutschen (nach Eisenberg 2013: 63).	24
Abbildung 2: Kompositionale Bedeutung prototypischer adverbialer PP und Objekt-PP.....	26
Abbildung 3: Bedeutungsspektrum von Präpositionen (P) in Objekt-PP (nach Breindl 1989: 42).	27
Abbildung 4: Semantisierte und desemantisierte Präpositionen in Objekt-PP (nach Breindl 1989: 39).....	28
Abbildung 5: Die Lernerkorpusforschung und ihre Schnittstellen: Überblick.	62
Abbildung 6: Ausgewählte Lerner- und Aufgabenvariablen (Metadaten) zur Dokumentation in einem Lernerkorpus (nach Granger 2008: 264).	65
Abbildung 7: Lernerkorpus-Typologie (nach Granger 2008 und Gilquin 2015).....	66
Abbildung 8: Beispiel für eine Zielhypothese (ZH). Entnommen aus dem Lernerkorpus MERLIN (ID: 1023_0001416).	73
Abbildung 9: Beispiel für die ZH-Varianten ZH0, ZH1 und ZH2 am Beispiel einer Lerneräußerung aus MERLIN (ID: 1023_0001418).....	73
Abbildung 10: Beispiel für die Rekonstruktion eines ausgelassenen Elements (hier: Präposition auf) auf Ebene der ZH. Entnommen aus dem Lernerkorpus MERLIN (ID: 1071_0024874). O=Original-Lerneräußerung.	74
Abbildung 11: Beispiel für die Editier-Tags INS (repräsentiert eine Einfügung) und CHA (repräsentiert eine Änderung). Entnommen aus dem Lernerkorpus MERLIN (ID: 1071_0024874).....	75
Abbildung 12: Beispiel für das Editier-Tag DEL (repräsentiert eine Löschung). Entnommen aus dem Lernerkorpus MERLIN (ID: 1061_0120338).....	76
Abbildung 13: Die traditionelle Fehleranalyse und die computerunterstützte Fehleranalyse: Gegenüberstellung.....	82

Abbildung 14: Kontrastive Interlanguage-Analyse (contrastive interlanguage analysis, CIA) (nach Granger 2015: 17).	85
Abbildung 15: Grundlagen der Dissertation: Zusammenfassung.....	124
Abbildung 16: Forschungsstand zum Gebrauch von Präpositionen und PP durch Deutschlernende: Zusammenfassung.	126
Abbildung 17: Inhaltliche und methodische Ausrichtung der Dissertation.....	127
Abbildung 18: Datenerhebungsdesign longitudinal (L) und quasi-longitudinal (QL): Vergleich (nach Meunier 2015: 382).....	129
Abbildung 19: Einordnung des Lernerkorpus MERLIN in die Lernerkorpus-Typologie.	130
Abbildung 20: Kompetenzniveaus gemäß dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GeR) (Europarat 2001: 41).	131
Abbildung 21: Aufgabenstellung GeR-Testniveau A2: Beispiel (Ausschnitt).	131
Abbildung 22: Aufgabenstellung GeR-Testniveau C1: Beispiel (Ausschnitt).	132
Abbildung 23: Verteilung der Lernertexte in MERLIN je GeR-Testniveau und GeR-Gesamtniveau: Vergleich (prozentual).	134
Abbildung 24: Beispiel für automatische Annotationen in MERLIN: Part-of-Speech (POS) (oben) und Dependenzstrukturen (unten).	136
Abbildung 25: Die Ebene ZH1Diff in MERLIN: Beispiel 1. O=Original-Lerneräußerung.....	138
Abbildung 26: Die Ebene ZH1Diff in MERLIN: Beispiel 2. O=Original-Lerneräußerung.....	138
Abbildung 27: Aufbau von Fehlertags in MERLIN: Beispiel.....	140
Abbildung 28: Überblick über untersuchte Einflussvariablen im Gebrauch von Präpositionen und PP durch Deutschlernende.....	143
Abbildung 29: Output eines Regressionsmodells mit zwei unabhängigen Variablen und einer Interaktion zwischen den Variablen: Beispiel aus Gries (2013: 301).	148
Abbildung 30: Output eines Regressionsmodells mit zwei kategorialen Variablen (SPEAKER und GeR) mit unterschiedlicher Anzahl an Ausprägungen: Beispiel.....	151
Abbildung 31: Output der Anova()-Funktion in R zu einem Regressionsmodell mit zwei unabhängigen kategorialen Variablen: Beispiel.....	152
Abbildung 32: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM1. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.	166
Abbildung 33: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM1a.	168
Abbildung 34: Fehlertyp Wahl im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM2. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.....	174
Abbildung 35: Fehlertyp Wahl im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM2a.	175
Abbildung 36: Fehlertyp Realisierung im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Interaktion): Ergebnis des Regressionsmodells RM3. L1TV = gestrichelte Linie, L1TA = durchgezogene Linie.....	178
Abbildung 37: Fehlertyp Realisierung im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM3a.	179
Abbildung 38: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM1. Objekt-PP = gestrichelte Linie, adverbiale PP = durchgezogene Linie.	202

Abbildung 39: Fehlertypen Wahl (links) und Auslassung (rechts) im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau: Ergebnisse der Regressionsmodelle RM2 und RM3. Objekt-PP = gestrichelte Linie, adverbiale PP = durchgezogene Linie.	208
Abbildung 40: Fehlertyp Auslassung je syntaktischer Funktion der PP und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM2a. A = adverbiale PP, O = Objekt-PP.	210
Abbildung 41: Fehlertyp Wahl je syntaktischer Funktion der PP und GeR-Gesamtniveaus (Einzeleffekte). Ergebnis des Regressionsmodells RM3a. A = adverbiale PP, O = Objekt-PP.	211
Abbildung 42: Korrektheit im Präpositionsgebrauch Präposition (semP/desemP) und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM1, desemP = durchgezogene Linie, semP = gestrichelte Linie.....	235
Abbildung 43: Korrektheit im Präpositionsgebrauch in Objekt-PP Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte): Ergebnis des Regressionsmodells RM1a.....	236
Abbildung 44: Fehlertyp Hinzufügung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau: Ergebnis des Regressionsmodells RM2, desemP = durchgezogene Linie, semP = gestrichelte Linie.....	242
Abbildung 45: Fehlertyp Hinzufügung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM2a.	243
Abbildung 46: Fehlertyp Hinzufügung im Präpositionsgebrauch in Objekt-PP Präposition (semP/desemP) (Einzeleffekt). Ergebnis des Regressionsmodells RM2b.....	244
Abbildung 47: Fehlertyp Auslassung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau: Ergebnisse der Regressionsmodelle RM3, desemP = durchgezogene Linie, semP = gestrichelte Linie.....	247
Abbildung 48: Fehlertyp Auslassung im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM3a.	248
Abbildung 49: Fehlertyp Wahl im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau. Ergebnis des Regressionsmodells RM4, desemP = durchgezogene Linie, semP = gestrichelte Linie.....	249
Abbildung 50: Fehlertyp Wahl im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau (Einzeleffekte). Ergebnis des Regressionsmodells RM4a.	250

Tabellenverzeichnis

Tabelle 1: Übersicht zur Kasusreaktion von Präpositionen.	8
Tabelle 2: Semantischer Gehalt von Präpositionen: Überblick.....	14
Tabelle 3: Präpositionen: Zentrale Terminologie.	17
Tabelle 4: Syntaktische Funktionen von PP: Überblick.....	19
Tabelle 5: Objekt-PP und adverbiale PP: Form-Funktionsverbindungen.	25
Tabelle 6: Semantisierte und desemantisierte Präpositionen in Objekt-PP: Form-Funktionsverbindungen.	30
Tabelle 7: Präpositionalphrasen aus syntaktischer Perspektive: Zentrale Terminologie.....	33
Tabelle 8: Argumentstatus von PP (nach Primus 2012: 4).....	37
Tabelle 9: Ähnliche semantische Rollen bei Objekt-PP und adverbialen PP: Beispiele (nach Breindl 1989: 54f.).	38
Tabelle 10: Präpositionalphrasen unter valenzgrammatischer Perspektive: Zentrale Terminologie.....	41

Tabelle 11: Andere Verwendungsweisen von Präpositionalphrasen: Zentrale Terminologie.	45
Tabelle 12: PP-Äquivalente im Türkischen und Ungarischen: Beispiele.....	47
Tabelle 13: PP-Äquivalente im Englischen und Italienischen: Beispiele.....	48
Tabelle 14: Präpositionalphrasen: Zusammenfassung der verschiedenen Perspektiven und Analyseebenen.	49
Tabelle 15: Ausgewählte Publikationen (deutsche Lernerkorpora) an der Schnittstelle zur Zweit- und Fremdspracherwerbsforschung.	98
Tabelle 16: Untersuchungen zu Präpositionen und PP im DaZ/DaF-Erwerb: Überblick.	102
Tabelle 17: Forschungsstand zu ausgewählten Einflussvariablen im Gebrauch von Präpositionen und PP in der Zielsprache Deutsch: Zusammenfassung.....	117
Tabelle 18: Verteilung der Lernertexte und Token in MERLIN je GeR-Testniveau: Übersicht.	132
Tabelle 19: GeR-Testniveau und GeR-Gesamtniveau: Gegenüberstellung.	133
Tabelle 20: Verteilung der Lernertexte und Token in MERLIN je GeR-Gesamtniveau: Überblick.	133
Tabelle 21: Metadaten in MERLIN: Überblick.	135
Tabelle 22: Zentrale automatische Annotationen in MERLIN: Überblick.....	136
Tabelle 23: Zentrale manuelle Annotationen in MERLIN: Überblick.....	137
Tabelle 24: Fehlerannotationskategorien und Fehlertags im Bereich Präpositionsgebrauch in MERLIN (Wisniewski et al. 2014: 12).	140
Tabelle 25: Fehlerannotationskategorien im Präpositionsgebrauch in MERLIN: Beispiele. .	145
Tabelle 26: Analysemethoden der Fallstudien: Überblick.	146
Tabelle 27: Übersicht zu untersuchten Erstsprachen (L1) der Lernenden.	155
Tabelle 28: Zusammenstellung der Subkorpora L1TV und L1TA in MERLIN: Überblick.....	157
Tabelle 29: Korpusabfragen zur Fallstudie Erstsprache der Lernenden.	158
Tabelle 30: Ergebnis der Korpusabfragen je L1-Subkorpus (nicht bereinigt).	159
Tabelle 31: PP-Kontexte je L1-Subkorpus (nicht bereinigt).	159
Tabelle 32: PP-Kontexte je L1-Subkorpus (bereinigt).	160
Tabelle 33: Verteilung der PP-Kontexte auf unterschiedliche Lernertexte je L1-Subkorpus. AL=Anzahl Lernertexte.	161
Tabelle 34: Analysemethoden der Fallstudie zur Variable Erstsprache der Lernenden: Überblick.	161
Tabelle 35: Korrektheit im Präpositionsgebrauch je L1-Subkorpus: Überblick.	162
Tabelle 36: Korrektheit im Präpositionsgebrauch je L1-Subkorpus und GeR-Gesamtniveau.	163
Tabelle 37: PP-Kontexte je Erstsprache und GeR-Gesamtniveau: Überblick.....	164
Tabelle 38: Korrekte und inkorrekte PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau: Überblick. AL=Anzahl unterschiedlicher Lernertexte.	164
Tabelle 39: Inkorrekte PP-Kontexte je L1-Subkorpus und GeR-Gesamtniveau: Überblick. AL=Anzahl Lernertexte.	171
Tabelle 40: Fehlertypen im Präpositionsgebrauch je L1-Subkorpus (GeR-Gesamtniveaus A2, B1+ und B2): Überblick.	171
Tabelle 41: Fehlertypen Wahl (W) und Realisierung (R) je L1-Subkorpus und GeR-Gesamtniveau.....	173
Tabelle 42: Objekt-PP und adverbiale PP: Form-Funktionsverbindungen.	186
Tabelle 43: Die sechs häufigsten Präpositionen (inkl. Verschmelzungen) in dlexDB.....	188
Tabelle 44: Korpusabfragen zur Fallstudie Syntaktische Funktion der PP.....	190

Tabelle 45: Ergebnis der Korpusabfragen je Präposition (Wortform) und GeR-Gesamtniveau (nicht bereinigt).	190
Tabelle 46: Annotationskategorien in der Fallstudie Syntaktische Funktion der PP.	192
Tabelle 47: Annotationskategorien und -werte zum Fehlertyp im Bereich Präpositionsgebrauch in der Fallstudie Syntaktische Funktion der PP.	193
Tabelle 48: PP-Kontexte je Präposition (an/auf), GeR-Gesamtniveau und syntaktischer Funktion der PP (bereinigt).	197
Tabelle 49: PP-Kontexte je GeR-Gesamtniveau und syntaktischer Funktion der PP.	198
Tabelle 50: Anzahl unterschiedlicher Lernertexte je GeR-Gesamtniveau: Gesamtverteilung und getrennt nach syntaktischer Funktion der PP.	198
Tabelle 51: Analysemethoden der Fallstudie zur Variable Syntaktische Funktion der PP: Überblick.	199
Tabelle 52: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP: Überblick.	200
Tabelle 53: Korrektheit im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR- Gesamtniveau. AL=Anzahl Lernertexte.....	200
Tabelle 54: Inkorrekte PP-Kontexte je syntaktischer Funktion der PP und GeR-Gesamtniveau: Überblick. AL=Anzahl Lernertexte.	205
Tabelle 55: Fehlertypen im Präpositionsgebrauch je syntaktischer Funktion (GeR- Gesamtniveaus A2, B1, B2 und C1): Überblick.....	205
Tabelle 56: Fehlertypen Wahl und Auslassung im Präpositionsgebrauch je syntaktischer Funktion der PP und GeR-Gesamtniveau.....	206
Tabelle 57: Semantisierte und desemantisierte Präpositionen (in Objekt-PP): Form- Funktionsverbindungen.	219
Tabelle 58: Korpusabfragen zur Fallstudie Semantischer Gehalt der Präposition.....	222
Tabelle 59: Ergebnis der Korpusabfragen zur Fallstudie Semantischer Gehalt der Präposition (Wortform) je Präposition und GeR-Gesamtniveau (nicht bereinigt).	222
Tabelle 60: Annotationskategorien in der Fallstudie Semantischer Gehalt der Präposition.	225
Tabelle 61: Annotationskategorien und -werte zum Fehlertyp im Bereich Präpositionsgebrauch in der Fallstudie Semantischer Gehalt der Präposition.	226
Tabelle 62: PP-Kontexte je Präposition (semP/desemP) und GeR-Gesamtniveau (bereinigt).	230
Tabelle 63: Anzahl unterschiedlicher Lernertexte je GeR-Gesamtniveau: Gesamtverteilung und getrennt nach semantischem Gehalt der Präposition.....	231
Tabelle 64: Analysemethoden der Fallstudie zur Variable Semantischer Gehalt der Präposition: Überblick.....	232
Tabelle 65: Korrektheit im Präpositionsgebrauch in Objekt-PP je Präposition (semP/desemP): Überblick.	233
Tabelle 66: Korrektheit im Präpositionsgebrauch in Objekt-PP Präposition und GeR- Gesamtniveau. AL=Anzahl Lernertexte.....	233
Tabelle 67: Inkorrekte PP-Kontexte je Präposition (semP/desemP) in Objekt-PP und GeR- Gesamtniveau. AL= Anzahl Lernertexte.....	238
Tabelle 68: Fehlertypen in Objekt-PP je Präposition (semP/desemP) (GeR-Gesamtniveaus A2, B1, B2 und C1): Überblick.....	239
Tabelle 69: Fehlertypen Wahl, Auslassung und Hinzufügung in Objekt-PP je Präposition (semP/desemP) und GeR-Gesamtniveau.....	241
Tabelle 70: Fallstudien im Lernerkorpus MERLIN: Zentrale Ergebnisse zur Korrektheit und Fehlertypen im Präpositionsgebrauch.	258

Literaturverzeichnis

- Abel, Andrea (2010): Sprachtests und soziale Implikationen. In: *Deutsch als Fremdsprache* 4, S. 202–209.
- Abel, Andrea/Glaznieks, Aivars/Nicolas, Lionel/Stemle, Egon (2014): KoKo: An L1 Learner Corpus for German. In: Calzolari, Nicoletta/Choukri, Khalid/Declerck, Thierry/Loftsson, Hrafn/Maegaard, Bente/Mariani, Joseph/Moreno, Asuncion/Odijk, Jan/Piperidis, Stelios (Hrg.): *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavik, 26-31 May, 2014. Reykjavik: European Language Resource Association. S. 2414-2421.
- Abel, Andrea/Glaznieks, Aivars/Nicolas, Lionel/Stemle, Egon (2016): An extended version of the KoKo German L1 Learner corpus. In: Corazza, Anna/Montemagni, Simonetta/Semeraro, Giovanni (Hrg.): *Proceedings of the Third Italian Conference on Computational Linguistics CLiC-it 2016*, 5-6 December 2016, Napoli. Torino: Accademia University Press. S. 13-18. DOI: 10.4000/books.aaccademia.1666 (Mai 2019).
- Abel, Andrea/Wisniewski, Katrin/Nicolas, Lionel/Boyd, Adriane/Hana, Jirka/Meurers, Detmar (2014): A Trilingual Learner Corpus Illustrating European Reference Levels. In: *Ricognizioni – Rivista di Lingue, Letterature e Culture Moderne* 2 (1), S. 111-126. DOI: <http://dx.doi.org/10.13135/2384-8987/702> (Mai 2019).
- Ágel, Vilmos/Eichinger, Ludwig/Eroms, Hans-Werner/Hellwig, Peter/Heringer, Hans-Jürgen/Lobin, Henning (Hrg.) (2003): *Dependenz und Valenz. Ein internationales Handbuch zeitgenössischer Forschung/Dependency and Valency. An International Handbook of Contemporary Research*. 1. Halbband. Berlin u.a.: de Gruyter. (= Handbücher zur Sprach- und Kommunikationswissenschaft)
- Ágel, Vilmos/Eichinger, Ludwig/ Eroms, Hans-Werner/Hellwig, Peter/Heringer, Hans-Jürgen/Lobin, Henning (Hrg.) (2006): *Dependenz und Valenz. Ein internationales Handbuch zeitgenössischer Forschung/Dependency and Valency. An International Handbook of Contemporary Research*. 2. Halbband. Berlin u.a.: de Gruyter. (= Handbücher zur Sprach- und Kommunikationswissenschaft).
- Ahrenholz, Bernt (2014): *Lernersprachenanalyse*. In: Settinieri, Julia/Demirkaya, Sevilen/Feldmeier, Alexis/Gültekin-Karakoç, Nazan/Riemer, Claudia (Hrg.): *Empirische Forschungsmethoden für Deutsch als Fremd- und Zweitsprache. Eine Einführung*. Paderborn: Ferdinand Schöningh. S. 167-181.
- Apeltauer, Ernst (1987): *Indikatoren zur Sprachstandsbestimmung ausländischer Schulanfänger*. In: Apeltauer, Ernst (Hrg.): *Gesteuerter Zweitspracherwerb. Voraussetzungen und Konsequenzen für den Unterricht*. München: Max Hueber. S. 207-232.
- Apeltauer, Ernst (2010): *Lernersprache(n)*. In: Krumm et al. (Hrg.), S. 833-842.
- Artstein, Ron/Poesio, Massimo (2008): *Inter-Coder Agreement for Computational Linguistics*. *Computational Linguistics* 34 (4), S. 555-596.
- Backhaus, Klaus/Erichson, Bernd/Plinke, Wulff/Weiber, Rolf (2016): *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin/Heidelberg: Springer. DOI: 10.1007/978-3-662-46076-4_6.
- Balci, Tahir (2006): *Valenzstrukturabhängige Probleme beim DaF-Lernen türkischer Studierender*. In: *Deutsch als Fremdsprache* 4, S. 239-241.
- Balci, Tahir/Kanatlı, Faik (2001): *Das Problem der Kasuswahl nach Wechselpräpositionen*. In: *Deutsch als Fremdsprache* 1, S. 28-30.
- Baten, Kristof (2008): *Der Regelkomplex der Wechselpräpositionen mit Blick auf den DaF-Unterricht*. In: *Deutsch als Fremdsprache* 1, S. 22-26.

- Baten, Kristof (2009). Wechselpräpositionen im DaF-Unterricht. In: Deutsch als Fremdsprache 2, S. 96-104.
- Baten, Kristof/Willems, Klaas (2012): Kasuserwerb in der Präpositionalphrase vom Standpunkt der Verarbeitbarkeitstheorie (Processability Theory). In: Deutsche Sprache 40, S. 221-239.
- Bates, Douglas/Maechler, Martin/Bolker, Ben/Walker, Steve (2015): Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67 (1), S. 1-48.
- Barker, Fiona/Salamoura, Angeliki/Saville, Nick (2015): Learner corpora and language testing. In: Granger/Gilquin/Meunier (Hrg.), S. 511-533.
- Becker, Angelika (2012): Konzeptorientierte Ansätze. Der Ausdruck von Raum. In: Ahrenholz, Bernd (Hrg.): Einblicke in die Zweitspracherwerbsforschung und ihre methodischen Verfahren. Berlin u.a.: de Gruyter. S. 27-48.
- Becker, Tabea (2006): Erwerb und Verarbeitung komplexer grammatischer Strukturen bei Grundschulkindern. Ein Vergleich zwischen ein- und zweisprachigen Kindern. In: Becker, Tabea/Peschel, Corinna (Hrg.): Gesteuerter und ungesteuerter Grammatikerwerb. Baltmannsweiler: Schneider Verlag Hohengehren. S. 156-173.
- Behrens, Heike (2009a): Usage-based and emergentist approaches to language acquisition. In: Linguistics 47 (2), S. 383-411.
- Behrens, Heike (2009b): Konstruktionen im Spracherwerb. In: Zeitschrift für Germanistische Linguistik 37 (3), S. 427-444.
- Bley-Vroman, Robert (1983): The comparative fallacy in interlanguage studies: The case of systematicity. In: Language Learning 33, S. 1-17.
- Blumenthal-Dramé, Alice/Kortmann, Bernd (2013): Die Verschiedenheit der Sprachen. In: Auer, Peter (Hrg.): Sprachwissenschaft. Grammatik – Interaktion – Kognition. Weimar: J.B. Metzler. S. 285-317.
- Bouillon, Henri (1977): Einiges zu den Präpositionsfehlern im Fremdsprachenunterricht In: Germanistische Mitteilungen 5, S. 37-66.
- Brdar-Szabó, Rita (2010a): Nutzen und Grenzen der kontrastiven Analyse für Deutsch als Fremd- und Zweitsprache. In: Krumm et al. (Hrg.), S. 518-531.
- Brdar-Szabó, Rita (2010b): Kontrastive Analyse Ungarisch-Deutsch. In: Krumm et al. (Hrg.), S. 732-737.
- Breckle, Margit (2014): Zur korpusbasierten Analyse von Pronomina in DaF-Lernertexten: Eine Pilotstudie. In: Lamminpää, Siru/Rink, Christian (Hrg.): Demokratia, Demokrati, Democracy, Demokratie. VAKKI-Symposium XXXIV 13.-14.2.2014. Vaasa, S. 73-85. http://www.vakki.net/publications/no3_fin.html (Mai 2019).
- Breckle, Margit/Zinsmeister, Heike (2010): Zur lernersprachlichen Generierung referierender Ausdrücke in argumentativen Texten. In: Skiba, Dirk (Hrg.): Textmuster: schulisch - universitär – kulturkontrastiv: Beiträge zur Theorie und Empirie lernersprachlicher Textproduktion. Frankfurt a.M. u.a.: Lang. S. 79-101.
- Breckle, Margit/Zinsmeister, Heike (2012): A corpus-based contrastive analysis of local coherence in L1 and L2 German. In: Karabalic, Vladimir/Varga, Melita Aleksa/Pon, Leonard (Hrg.): Discourse and Dialogue/Diskurs und Dialog. S. 235-250.
- Breckle, Margit/Zinsmeister, Heike (2013): L1 transfer versus fixed chunks: A learner corpus-based study of L2 German. In: Twenty Years of Learner Corpus Research: Looking back, Moving ahead. Corpora and Language in Use – Proceedings 1, Louvain-la-Neuve: Presses universitaires de Louvain. S. 25-35.
- Breindl, Eva (1989): Präpositionalobjekte und Präpositionalobjektsätze im Deutschen. Tübingen: Niemeyer. (= Linguistische Arbeiten 220).

- Breindl, Eva (2006): Präpositionalphrasen. In: Ágel et al. (Hrg.). S. 936-951.
- Breindl, Eva (2013): Komplexe grammatische Strukturen im L2-Erwerb: Das Beispiel der Verschmelzung von Präposition und Artikel. In: *German as a foreign language 2*, S. 1-20.
- Breindl, Eva (2016): Konnexion in argumentativen Texten. Gebrauchsunterschiede in Deutsch als L2 vs. Deutsch als L1. In: d'Avis, Franz/Lohnstein, Horst (2016): *Normalität in der Sprache*. Hamburg: Buske. S. 37-64. (= Linguistische Berichte Sonderheft 22).
- Breindl, Eva (2018): Konnexion in argumentativen Texten von DaF-Lernern und Muttersprachlern. In: *Deutsche Sprache 1*, S. 22-36.
- Burger, Harald (2015): *Phraseologie. Eine Einführung am Beispiel des Deutschen*. 5., neu bearbeitete Ausgabe. Berlin: Erich Schmidt. (=Grundlagen der Germanistik 36)
- Bryant, Doreen (2012): Lokalisierungsausdrücke im Erst- und Zweitspracherwerb. Typologische, ontogenetische und kognitionspsychologische Überlegungen zur Sprachförderung in DaZ. Baltmannsweiler: Schneider Verlag Hohengehren.
- Cahill, Aoife/Madnani, Nitin/Tetreault, Joel R./Napolitano, Diane (2013): Robust systems for preposition error correction using Wikipedia revisions. In: *Proceedings of NAACL-HLT 2013. The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 9 - 14 June 2013. S. 507-517.
- Callies, Marcus (2015): Learner corpus methodology. In: Granger/Gilquin/Meunier (Hrg.), S. 35-55.
- Carletta, Jean (1996): Squibs and Discussions. Assessing Agreement on Classification Tasks: The Kappa Statistic. In: *Computational Linguistics 22 (2)*, S. 249-254.
- Carletta, Jean/Evert, Stefan/Heid, Ulrich/Kilgour, Jonathan/Robertson, Judy (2003): The NITE XML Toolkit: Flexible annotation for multimodal language data. In: *Behavior Research Methods, Instrument, & Computers 25 (3)*, S. 353-363.
- Carlsen, Cecile (2012): Proficiency Level – a Fuzzy Variable in Computer Learner Corpora. In: *Applied Linguistics 33 (2)*, S. 161-181.
- Carroll, Mary/Becker, Angelika u.a. (1993): Reference to space in learner varieties. In: Perdue, Clive (Hrg.): *Adult language acquisition: cross-linguistic perspectives. Volume II. The results*. Cambridge: Cambridge UP. S. 119-149.
- Chambers, Angela (2015): The learner corpus as a pedagogic corpus. In: Granger/Gilquin/Meunier (Hrg.), S. 445-484.
- Cherubim, Dieter (1980) (Hrg.): *Fehlerlinguistik. Beiträge zum Problem der sprachlichen Abweichung*. Tübingen: Niemeyer.
- Chodorow, Martin/Tetreault, Joel/Han, Na-Rae (2007): Detection of Grammatical Errors Involving Prepositions. In: *Proceedings of the 4th ACL-SIGSEM Workshop on Prepositions*. Prag, Tschechien. Juni 2007. S. 25-30.
- Christ, Oliver (1994): A Modular and Flexible Architecture for an Integrated Corpus Query System. In: Kiefer, Ferenc/Kiss, Gábor/Pajzs, Júlia (Hrg.): *Papers in Computational Lexicography, COMPLEX '94*. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences. S. 23-32.
- Clahsen, Harald/Meisel, Jürgen/Pienemann, Manfred (1983): *Deutsch als Zweitsprache - Der Spracherwerb ausländischer Arbeiter*. Tübingen: Narr.
- Clyne, Michael (1968): Zum Pidgin-Deutsch der Gastarbeiter. In: *Zeitschrift für Mundartforschung 35 (2)*, S. 130-139.
- Cohen, Jacob (1960): A coefficient of agreement for nominal scales. In: *Educational and psychological measurement 20 (1)*, S. 37-46.

- Corder, Stephen P. (1975): Error Analysis, Interlanguage and Second Language Acquisition. In: *Language Teaching* 8, S. 201-218.
- Costa, Marcella (2010): Kontrastive Analyse Italienisch-Deutsch. In: Krumm et al. (Hrg.), S. 586-593.
- Csillag, Enikő/Károlyi, Andrea/Nagy, Anna (1983): Die deutschen Präpositionen und ihre ungarischen Entsprechungen - Didaktische Möglichkeiten der Bekämpfung der Interferenz der Muttersprache. In: *Deutsch als Fremdsprache* 4, S. 230-234.
- Dagneaux, Estelle/Denness, Sharon/Granger, Sylviane (1998): Computer-aided error analysis. In: *System: An International Journal of Educational Technology and Applied Linguistics*, 26 (2), S. 163-174.
- DeKeyser, Robert M. (2003): Implicit and Implicit Learning. In: Doughty, Catherine J./Long, Michael H. (Hrg.): *The Handbook of Second Language Acquisition*. Malden: Blackwell. S. 313-348.
- DeKeyser, Robert M. (2005): What Makes Learning L2 Grammar difficult? In: *Language Learning* 55 (1), S. 1-25.
- dlexDB = lexikalische Datenbank. Universität Potsdam und BBAW. <http://www.dlexdb.de>.
- Díaz-Negrillo, Ana/Fernández-Domínguez, Jesús (2006): Error tagging systems for learner corpora. In: *Spanish Journal of Applied Linguistics (RESLA)* 19, S. 83-102.
- Díaz-Negrillo, Ana/Meurers, Detmar/Salvador, Valera/Wunsch, Holger (2010): Towards interlanguage POS annotation for effective learner corpora in SLA and FLT. In: *Language Forum* 36 (1-2), S. 139-154.
- Díaz-Negrillo, Ana/Thompson, Paul (2013): Learner Corpora. Looking towards the future. In: Díaz-Negrillo, Ana/Ballier, Nicolas/Thompson, Paul (Hrg.): *Automatic Treatment and Analysis of Learner Corpus Data*. Amsterdam u.a.: John Benjamins. S. 9-29.
- Dickinson, Markus/Ragheb, Marwa (2009): Dependency Annotation for Learner Corpora. In: Passarotti, Marco/Przepiórkowski, Adam/Raynaud, Savina/Van Eynde, Frank (Hrg.): *Proceedings of the Eighth International Workshop on Treebanks and Linguistic Theories 4-5 Dezember 2009 Mailand, Italien*. Mailand: EDUCatt. S. 59-70. http://convegni.unicatt.it/meetings_5737.html (Mai 2019).
- Diehl, Erika/Christen, Helen/Leuenberger, Sandra/Pelvat, Isabelle/Studer, Thérèse (2000): *Grammatikunterricht: Alles für die Katz? Untersuchungen zum Zweitspracherwerb Deutsch*. Tübingen: Niemeyer. (= Reihe Germanistische Linguistik 220).
- Doughty, Catherine J./Long, Michael H. (Hrg.) (2003): *The Handbook of Second Language Acquisition*. Malden: Blackwell.
- Duden (2016): *Duden. Die Grammatik. 9., vollständig überarbeitete und aktualisierte Auflage*. Herausgegeben von Angelika Wöllstein und der Dudenredaktion. Berlin: Dudenverlag. (= Band 14 – Der Duden in 12 Bänden).
- Duden-Online: Duden Online-Wörterbuch. Bibliographisches Institut Berlin. <http://www.duden.de/woerterbuch>.
- DWDS = DWDS Digitales Wörterbuch der deutschen Sprache (Berlin-Brandenburgische Akademie der Wissenschaften, Berlin). <http://www.dwds.de>.
- E-Valbu – Das elektronische Valenzwörterbuch deutscher Verben. Bis November 2018 unter folgender Adresse erreichbar: <http://hypermedia.ids-mannheim.de/evalbu/index.html> (Juli 2018); ab Dezember 2018 erreichbar unter: <https://grammis.ids-mannheim.de/verbvalenz> (Mai 2019).
- Eeg-Olofsson, Jens/Knutsson, Ola (2003): Automatic Grammar Checking for Second Language Learners - the Use of Prepositions. In: *Proceedings of the Nordic Conference of Computational Linguistics (NoDaLiDa)*, Reykjavik, Iceland.

- <http://www.nada.kth.se/~knutsson/publications.html> (Mai 2019).
- Eichinger, Ludwig/Kubczak, Jaqueline/Berens, Franz Josef (Hrg.) (2011): *Dependenz, Valenz und mehr: Beiträge zum 80. Geburtstag von Ulrich Engel*. Tübingen: Julius Groos.
- Eisenberg, Peter (2013): *Grundriss der deutschen Grammatik. Band 2: Der Satz. 4., aktualisierte und überarbeitete Auflage*. Stuttgart u.a.: J.B. Metzler.
- Ellis, Nick C. (2002): *Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition*. In: *Studies in Second Language Acquisition* 4, S. 143-188.
- Ellis, Nick C. (2003): *Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure*. In: Doughty, Catherine J./Long, Michael H. (Hrg.): *The Handbook of Second Language Acquisition*. Malden: Blackwell. S. 63-103.
- Ellis, Nick C. (2006a): *Language Acquisition as Rational Contingency Learning*. *Applied Linguistics* 2006 27 (1), S. 1-24.
- Ellis, Nick C. (2006b): *Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning*. In: *Applied Linguistics* 2006 27 (2), S. 164-194.
- Ellis, Nick C. (2008a): *The Associative-Cognitive CREED*. In: VanPatten, Bill/Williams, Jessica (Hrg.): *Theories in second language acquisition: An Introduction*. New York u.a.: Routledge. S. 77-95. (= Second language acquisition research theoretical and methodological issues).
- Ellis, Nick C. (2008b): *Usage-based and form-focused language acquisition*. In: Ellis, Nick C./Robinson, Peter (Hrg.): *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York u.a.: Routledge. S. 372-405.
- Ellis, Nick C./Wulff, Stefanie (2015): *Usage-based Approaches to SLA*. In: VanPatten, Bill/Williams, Jessica (Hrg.): *Theories in second language acquisition: An Introduction. Second Edition*. New York u.a.: Routledge. S. 75-93.
- Ellis, Rod (1994): *The Study of Second Language Acquisition*. Oxford u.a.: Oxford UP.
- Ellis, Rod/Barkhuizen, Gary (2005): *Analysing Learner Language*. Oxford/New York: Oxford University Press.
- Elsen, Hilke (2014): *Linguistische Theorien*. Tübingen: Narr.
- Engelberg, Stefan (2003): *"Glaube mich - ich habe es versucht."* Typische Lernerfehler und der Beitrag der neueren Valenzforschung zur Lernerlexikographie. In: Cyrus, Lea/Feddes, Hendrik/Schumacher, Frank/Steiner, Petra (Hrg.): *Sprache zwischen Theorie und Technologie*. Wiesbaden: Deutscher Universitäts-Verlag. S. 51-71.
- Engelberg, Stefan (2010): *Die lexikographische Behandlung von Argumentstrukturvarianten in Valenz- und Lernerwörterbüchern..* In: Schierholz, Stefan J. /Fischer, Klaus/Fobbe, Eilika (Hrg.): *Valenz und Deutsch als Fremdsprache*. Frankfurt/M.: Lang, S. 113-141.
- Engelberg, Stefan/König, Svenja/Proost, Kristel/Winkler, Edeltraud (2011): *Argumentstrukturmuster als Konstruktion? Identität – Verwandtschaft – Idiosynkrasien*. In: Engelberg, Stefan/Holler, Anke/Proost, Kristel (Hrg.): *Sprachliches Wissen zwischen Lexikon und Grammatik. Jahrbuch des Instituts für Deutsche Sprache 2010*. Berlin/Boston: de Gruyter. S. 71-112.
- Eroms, Hans-Werner (1991): *Valenzgebundene Präpositionalkonstruktionen im Deutschen*. In: Brinkmann, Hennig/Harweg, Roland/Kishitani, Shōko/Scherner, Maximilian (Hrg.): *Die deutsche Sprache – Gestalt und Leistung. H. Brinkmann in der Diskussion; zu seinem 90. Geburtstag*. Münster: Nodus. S. 39-54.

- Europarat (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Übersetzung: Quetz, Jürgen (in Zusammenarbeit mit Schieß, Raimund und Sköries, Ulrike)/Schneider, Günther. Berlin u.a.: Langenscheidt.
- Fandrych, Christian/Thurmair, Maria (2018): Grammatik im Fach Deutsch als Fremd- und Zweitsprache. Grundlagen und Vermittlung. Berlin: Erich Schmidt.
- Fandrych, Christian/Tschirner, Erwin (2007): Korpuslinguistik und Deutsch als Fremdsprache: Ein Perspektivenwechsel. In: Deutsch als Fremdsprache 4, S. 195-204.
- Fekete, Olga (2008): Forschungsmethodologische Aspekte zur Kasusverwendung bei ungarischen DaF-Lernenden. In: Böttger, Lydia/Masáth, András (Hrg.): Jahrbuch der Ungarischen Germanistik, Budapest: Gondolat Kiadó Kör Budapest. S. 163-183.
- Fischer, Kerstin/Stefanowitsch, Anatol (Hrg.) (2006/2007): Konstruktionsgrammatik. Von der Anwendung zur Theorie. Tübingen: Stauffenberg.
- Fitschen, Arne/Gupta, Piklu (2008): Lemmatising and morphological tagging. In: Lüdeling/Kytö (Hrg.) (2008), S. 552-564.
- Forgács, Tamás (2004): Ungarische Grammatik. 2., verbesserte Auflage. Wien: EDITION PRAESENS. Verlag für Literatur- und Sprachwissenschaft.
- Fox, John (2003): Effect Displays in R for Generalised Linear Models. Journal of Statistical Software 8 (15), S. 1-27. <http://www.jstatsoft.org/v08/i15/>. (Mai 2019).
- Fox, John/Hong, Jangman (2009): Effect Displays in R for Multinomial and Proportional Odds Logit Models: Extensions to the effects Package. Journal of Statistical Software 32 (1), S. 1-24. <http://www.jstatsoft.org/v32/i01/>. (Mai 2019).
- Fox, John/Weisberg, Sanford (2011). An R Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion-2E/index.html>.
- Freitag, York/Vandermeeren, Sonja (2005): Deutsche Präpositionen – Eine fehleranalytische Untersuchung. In: Das Wort. Germanistisches Jahrbuch Russland, S. 155-181.
- Fries, Nobert (1988): Deutsche Präpositionen und Präpositionalphrasen aus Sicht der kontrastiven Linguistik. In: Deutsch als Fremdsprache 6, S. 331-336.
- Gamon, Michael/Chodorow, Martin/Leacock, Claudia/Tetreault, Joel (2013): Using learner corpora for automatic error detection and correction. In: Díaz-Negrillo, Ana/Ballier, Nicolas/Thompson, Paul (Hrg.): Automatic Treatment and Analysis of Learner Corpus Data. Amsterdam u.a.: John Benjamins. S. 127- 149.
- Gass, Susan M./Mackey, Alison (2002): Frequency Effects and Second Language Acquisition. A Complex Picture? In: Studies in Second Language Acquisition 4, S. 249-260.
- Gass, Susan M./Selinker, Larry (2008): Second Language Acquisition. An Introductory Course. New York u.a.: Routledge.
- GDS (1997) = Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno (1997): Grammatik der deutschen Sprache. Band 1-3. Berlin u.a.: de Gruyter.
- Geyken, Alexander (2007): The DWDS corpus: A reference corpus for the German language of the 20th century. In: Fellbaum, Christiane (Hrg.): Idioms and Collocations: Corpus-based Linguistic und Lexicographic Studies. London: Continuum. S. 23-41.
- Gibson, Martha/Hufeisen, Britta/Libben, Gary (2001): Learners of German as an L3 and their production of German prepositional verbs. In: Cenoz, Jasone/Hufeisen, Britta/Jessner, Ulrike (Hrg.): Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives. Clevedon: Multilingual Matters. S. 138-148. (= Bilingual Education and Bilingualism 31).
- Gilquin, Gaëtanelle (2015): From design to collection of learner corpora. In: Granger/Gilquin/Meunier (Hrg.), S. 9-34.

- Gilquin, Gaëtanelle/Granger, Sylviane (2010): How can data-driven learning be used in language teaching? In: O’Keeffe, Anne/McCarthy, Michael (Hrg.): The Routledge handbook of corpus linguistics. London: Routledge. S. 359-370.
- Gilquin, Gaëtanelle/Granger, Sylviane/Paquot (2007): Learner Corpora: The missing link in EAP pedagogy. In: Journal of English for Academic Purposes 6, S. 319-333.
- Gilquin, Gaëtanelle/Paquot, Magali (2008): Too chatty: Learner academic writing and register variation. In: English Text Construction 1 (1), S. 41-61.
- Glaboniat, Manuela (2010): Profile deutsch: gemeinsamer europäischer Referenzrahmen; Lernzielbestimmungen; Kannbeschreibungen; kommunikative Mittel; Niveau A1-A2, B1-B2; C1-C2; [CD-ROM Version 2.0 mit Begleitbuch]. Berlin u.a.: Langenscheidt.
- Glaznieks, Aivars/Abel, Andrea (2017): ‚So einen Fehler wird einem das ganze Leben lang verfolgen.‘ Empirische Untersuchung grammatischer Kompetenzen am Ende der Oberschule. In: Davies, Winifred V./Häcki-Buhofer, Annelies/Schmidlin, Regula/Wagner, Melanie/Wyss, Eva Lia (Hrg.): Standardsprache zwischen Norm und Praxis. Theoretische Betrachtungen, empirische Studien und sprachdidaktische Ausblicke. Tübingen: Narr. S. 235-278. (= Basler Studien zur deutschen Sprache und Literatur 99).
- Glück, Alexander/Leonardi, Mara Maya Victoria (i.Dr.): Zur Verwendung von Präpositionen in Texten und Diskursen von Südtiroler Maturanten. In: Habermann, Mechthild/Kürschner, Sebastian/Müller, Peter O. (Hrg.): Dialektale Daten: Erhebung – Aufbereitung – Auswertung.
- Goschler, Juliana/Stefanowitsch, Anatol (2014): Korpora in der Zweitspracherwerbsforschung: Sieben Probleme aus korpuslinguistischer Sicht. In: Ahrenholz, Bernt/Grommes, Patrick (Hrg.): Zweitspracherwerb im Jugendalter. Bern u.a.: de Gruyter. S. 341-360.
- Gottschein-Schramm, Barbara/Kalender, Susanne/Specht, Franz/Duckstein, Barbara (2017): Schritte Übungsgrammatik. Niveau A1-B1. Ismaning: Hueber.
- Götz, Sandra (2015): Tense and aspect errors in spoken learner language: Implications for language testing and assessment. In: Callies, Marcus (Hrg.): Learner Corpora in Language Testing and Assessment. Amsterdam: John Benjamins. S. 191-215. <https://www.academia.edu/10267526> (Mai 2019).
- grammis = Grammatisches Informationssystem. Leibniz-Institut für deutsche Sprache. <https://grammis.ids-mannheim.de/> (Mai 2019).
- Graën, Johannes/Schneider, Gerold (2017): Crossing the Border Twice. Reimporting Prepositions to Alleviate L1-Specific Transfer Errors. In: Volodina, Elena/Pilán, Ildikó/Broin, Lars/Grigonytė, Gintarė/Nilsson Björkenstam, Kristina (Hrg.): Proceedings of the Joint 6th Workshop on NLP for Computer Assisted Language Learning and 2nd Workshop on NLP for Research on Language Acquisition at NoDaLiDa, Gothenburg, 22 May 2017. Linköping Electronic Conference Proceedings. S. 18-26. (<http://www.ep.liu.se/ecp/134/ecp17134.pdf>) (Dezember 2019).
- Granger, Sylviane (1994): The Learner Corpus: a revolution in Applied Linguistics. In: English Today 39 (10), S. 25-33.
- Granger, Sylviane (1996): From CA to CIA and back: An Integrated Approach to Computerized Bilingual and Learner Corpora. In: Aijmer, Karin/Altenberg, Bengt/Johansson, Mats (Hrg.): Languages in Contrast. Text-based Cross-linguistic Studies. Lund: Lund University Press. S. 37-51
- Granger, Sylviane (1999): Use of tenses by advanced EFL learners: evidence from an error-tagged computer corpus. In: Hasselgård, Hilde/Oksefjell, Signe (Hrg.): Out of corpora. Studies in honour of Stig Johansson. Amsterdam & Atlanta: Rodopi. S. 191-202.

- Granger, Sylviane (2002): A Bird's Eye View on Learner Corpus Research: In: Granger, Sylviane/Hung, Joseph/Petch-Tyson, Stephanie (Hrg.): Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching. Amsterdam: John Benjamins. S. 3-33.
- Granger, Sylviane (2003): Error-tagged Learner Corpora and CALL: A Promising Synergy. In: CALICO Journal 20 (3), S. 465-480.
- Granger, Sylviane (2008): Learner Corpora. In: Lüdeling/Kytö (Hrg.): S. 259-275.
- Granger, Sylviane (2009): The contribution of learner corpora to second language acquisition and foreign language teaching. In: Aijmer, Karin (Hrg.): Corpora and Language Teaching. Amsterdam u.a.: John Benjamins. S. 13-32.
- Granger, Sylviane (2012): How to use Foreign and Second Language Learner Corpora. In: Gass, Susan M./Mackey, Alison (Hrg.): Research methods in second language acquisition. A practical guide. Chichester u.a.: Wiley-Blackwell. S. 7-29.
- Granger, Sylviane (2015a): Contrastive interlanguage analysis: A reappraisal. In: International Journal of Learner Corpus Research 1 (1), S. 7-24.
- Granger, Sylviane (2015b): The contribution of learner corpora to reference and instructional materials design. In: Granger/Gilquin/Meunier (Hrg.): S. 485-510.
- Granger, Sylviane/Dagneaux, Estelle/Meunier, Fanny/Paquot, Magali (2009): The International Corpus of Learner English. Version 2. Handbook and CD-Rom, Presses Universitaires de Louvain: Louvain-la-Neuve.
- Granger, Sylviane/Gilquin, Gaëtanelle/Meunier, Fanny (2015): Introduction: learner corpus research: past, present and future. In: Granger/Gilquin/Meunier (Hrg.), S. 1-6.
- Granger, Sylviane/Gilquin, Gaëtanelle/Meunier, Fanny (Hrg.) (2015): The Cambridge Handbook of Learner Corpus Research. Cambridge: Cambridge University Press.
- Grein, Marion (2013): Neurodidaktik. Grundlagen für Sprachlehrende. Ismaning: Hueber.
- Grießhaber, Wilhelm (1999a): Die relationierende Prozedur. Zu Grammatik und Pragmatik lokaler Präpositionen und ihrer Verwendung durch türkische Deutschlerner. Berlin u.a.: Waxmann. (= Mehrsprachigkeit 5).
- Grießhaber, Wilhelm (1999b): Präpositionen als relationierende Prozeduren. In: Redder, Angelika/Rehbein, Jochen (Hrg.): Grammatik und mentale Prozesse. Tübingen: Stauffenburg. S. 241-260.
- Grießhaber, Wilhelm (2007a): Präposition. In: Hoffmann, Ludger (Hrg.): Handbuch der deutschen Wortarten. Berlin u.a.: de Gruyter. S. 629-655.
- Grießhaber, Wilhelm (2007b): "und wir faren in die andere seite" - Der Gebrauch lokaler Präpositionen durch türkische Grundschüler. In: Meng, Katharina & Rehbein, Jochen (Hrg.) Kindliche Kommunikation - einsprachig und mehrsprachig. Münster u.a.: Waxmann. S. 371-392.
- Grießhaber, Wilhelm (2010): Spracherwerbsprozesse in Erst- und Zweitsprache. Duisburg: Universitätsverlag Rhein-Ruhr.
- Grießhaber, Wilhelm (2011): Präpositionen als relationierende Verfahren – Präpositionen vor dem Hintergrund des Türkischen. In: Jahrbuch Deutsch als Fremdsprache 37, S. 142-159.
- Gries, Stefan Th. (2013): Statistical tests for the analysis of learner corpus data. In: Díaz-Negrillo, Ana/Ballier, Nicolas/Thompson, Paul (Hrg.): Automatic Treatment and Analysis of Learner Corpus Data. Amsterdam u.a.: John Benjamins. S. 287-309.
- Gries, Stefan Th. (2015): Statistics for learner corpus research. In: Granger/Gilquin/Meunier (Hrg.), S. 159-181.
- Gries, Stefan Th./Wulff, Stefanie (2013): The genitive alternation in Chinese and German ESL

- learners: towards a multifactorial notion of context in learner corpus research. In: *International Journal of Corpus Linguistics* 18 (3), 327-356.
- Gürsoy, Erkan/Wilhelm, Nadine (2014): Präpositionen in Mathematik-Prüfungsaufgaben als spezifische Herausforderung für türkischsprachige Lernende mit Deutsch als Zweitsprache. In: Ahrenholz, Bernt/Grommes, Patrick (Hrg.): *Zweitspracherwerb im Jugendalter*. Bern u.a.: de Gruyter. S. 77-98.
- Gut, Ulrike (2012): The LeaP corpus. A multilingual corpus of spoken learner German and learner English. In: Schmidt Thomas/Wörner Kai (Hrg.): *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam: John Benjamins. S. 3-23.
- Gutzmann, Daniel/Turgay, Katharina (2011): Funktionale Kategorien in der PP und deren Zweitspracherwerb. In: *Zeitschrift für Sprachwissenschaft* 30, S. 169-221.
- Hall, Christopher (2010): Kontrastive Analyse Englisch-Deutsch. In: Krumm et al. (Hrg.), S. 550-561.
- Hall, Christopher/Newbrand, Denise/Ecke, Peter/Sperr, Ulrike/Marchand, Vanessa/Hayes, Lisa (2009): Learners' Implicit Assumptions About Syntactic Frames in New L3 Words: The Role of Cognates, Typological Proximity, and L2 Status. In: *Language Learning* 59, S. 153- 202.
- Hancke, Julia/Meurers, Detmar (2013): Exploring CEFR classification for German based on rich linguistic modeling. In: *Learner Corpus Research 2013. Book of Abstracts*. S. 54-56. Bergen, Norway. <http://purl.org/dm/papers/Hancke.Meurers-13.html> (Mai 2019).
- Harnisch, Ulrike (1993): Grammatische Progression – Ein alter Hut? Zur Zweisprachentwicklung türkischer Schulanfänger. In: *Deutsch lernen* 18, S. 313-334.
- Harrison, Julia (2015): The English Grammar Profile. In: Harrison, Julia/Barker, Fiona (Hrg.): *English Profile in Practice*. Cambridge: Cambridge University Press. S. 28-48.
- Hauschild, Alke (2014): *Praxis Grammatik Deutsch als Fremdsprache. Das große Lern- und Übungswerk. Mit extra Online-Übungen*. Stuttgart: PONS.
- Hawthorne, James (2005): The acquisition of prepositional constructions and their associated case-marking properties in the L2 German of L1 English speakers. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.501.1984&rep=rep1&type=pdf> (Mai 2019).
- Heidolph, Karl Erich/Flämig, Walter/Mosch, Wolfgang (1981): *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag.
- Heister, Julia/Würzner, Kay-Michael/Bubbenzer, Johannes/Pohl, Edmund/Hanneforth, Thomas/Geyken, Alexander/Kliegl, Reinhold (2011): dlexDB - eine lexikalische Datenbank für die psychologische und linguistische Forschung. In: *Psychologische Rundschau* 62 (1), S. 10-20.
- Helbig, Gerhard/Buscha, Joachim (2001a): *Leitfaden der deutschen Grammatik*. Berlin u.a.: Langenscheidt.
- Helbig, Gerhard/Buscha, Joachim (2001b): *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Berlin u.a.: Langenscheidt.
- Helbig, Gerhard/Götze, Lutz/Henrici, Gert/Krumm, Hans-Jürgen (Hrg.) (2001): *Deutsch als Fremdsprache. Ein internationales Handbuch*. Berlin/New York: de Gruyter. 1./2. Halbband (= Handbücher zur Sprach- und Kommunikationswissenschaft).
- Hepsöyler, Ender/Liebe-Harkort, Klaus (1991): *Muttersprache und Zweitsprache: Türkische Schulanfängerinnen und Schulanfänger in der Migration – Ein Vergleich*. Frankfurt a.M. u.a.: Lang.
- Higgins, Derrick/Ramineni, Chaitanya/Zechner, Klaus (2015): Learner corpora and automated scoring. In: Granger/Gilquin/Meunier (Hrg.), S. 587-604.

- Hirschmann, Hagen/Lüdeling, Anke/Rehbein, Ines/Reznicek, Marc/Zeldes, Amir (2013): Underuse of Syntactic Categories in Falko. A Case Study on Modification. In: Granger, Sylviane/Gilquin, Gaëtanelle /Meunier, Fanny (Hrg.): 20 years of learner corpus research. Looking back, Moving ahead. Corpora and Language in Use. Proceedings of the first Learner Corpus Research Conference (LCR 2011). Louvain la Neuve: Presses universitaires de Louvain: Louvain la Neuve. S. 223-235.
- Hirschmann, Hagen/Nolda, Andreas (2019): Dulko – auf dem Weg zu einem deutsch - ungarischen Lernerkorpus. In: Eichinger, Ludwig/Plewnia, Albrecht (Hrg.): Neues vom heutigen Deutsch. Empirisch – methodisch – theoretisch. Jahrbuch des Instituts für Deutsche Sprache 2018. Berlin u.a.: de Gruyter. S. 339-342.
- Hoffmann, Ludger (2016): Deutsche Grammatik: Grundlagen für Lehrerausbildung, Schule, Deutsch als Zweitsprache und Deutsch als Fremdsprache. 3., neu bearbeitete und erweiterte Auflage. Berlin: Erich Schmidt Verlag.
- HPD (1977): Heidelberger Forschungsprojekt „Pidgin-Deutsch spanischer und italienischer Arbeiter in der Bundesrepublik“: Die ungesteuerte Erlernung des Deutschen durch spanische und italienische Arbeiter; eine soziolinguistische Untersuchung. Osnabrück: Universität (= OBST Beihefte 2).
- Huddleston, Rodney/Pullum, Geoffrey K. u.a. (2006): The Cambridge Grammar of the English Language. 4th Edition. Cambridge u.a.: UP.
- Hufeisen, Britta/Gibson, Martha (2002): Production of Locative Prepositions by Learner of German as a Second Language. In: Barkowski, Hans/Faistauer, Renate (Hrg.): ... in Sachen Deutsch als Fremdsprache. Baltmannsweiler: Schneider Verlag Hohengehren. S. 73-90.
- Hufeisen, Britta/Riemer, Claudia (2010): Spracherwerb und Sprachenlernen: Modelle und theoretische Ansätze. In: Krumm et al. (Hrg.), S. 738-754.
- Hundt, Markus (2001): Grammatikalisierungsphänomene bei Präpositionalobjekten in der deutschen Sprache. In: Zeitschrift für germanistische Linguistik 29 (2), S. 167-191.
- I6doc (2019): International Corpus of Learner English V2. I6doc.com. The place for scientific books (<https://www.i6doc.com/en/book/?gcoi=28001100825990>) (Mai 2019)
- Jacobs, Joachim (2003): Die Problematik der Valenzebenen. In: HSK Valenz und Dependenz, S. 378-399.
- José-Domínguez-Vázquez, María/Eichinger, Ludwig (Hrg.) (2015): Valenz im Fokus: grammatische und lexikografische Studien. Festschrift für Jacqueline Kubczak. Mannheim: Institut für deutsche Sprache.
- Kalkavan-Aydin, Zeynep (2017): Lokale Präpositionen im Deutschen und ihre Entsprechungen im Türkischen im Fokus kindlicher Spracherwerbsprozesse. In: Ekinci, Yüskel/Montanari, Elke/Selmani, Lirim (Hrg.): Grammatik und Variation. Festschrift für Ludger Hoffmann zum 65. Geburtstag. Heidelberg: Synchron. S. 375-387.
- Kaltenbacher, Erika/Klages, Hana (2006): Sprachprofil und Sprachförderung bei Vorschulkindern mit Migrationshintergrund. In: Ahrenholz, Bernt (Hrg.): Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten. Freiburg: Fillibach. S. 80-97.
- Keibel, Holger/Perkuhn, Rainer/Kupietz, Marc (2012): Korpuslinguistik. Paderborn: Wilhelm Fink.
- Keim, Inken (1984): Untersuchungen zum Deutsch türkischer Arbeiter. Tübingen: Narr.
- Kempe, Vera/MacWhinney, Brian (1998): The acquisition of case marking by adult learners of Russian and German. In: Studies in Second Language Acquisition 20, S. 543-587.

- Kidd, Evan/Cameron-Faulkner, Thea (2008): The acquisition of the multiple senses of with. In: *Linguistics* 46, S. 33-62.
- Krause, Thomas/Zeldes, Amir (2016): ANNIS3: A new architecture for generic corpus query and visualization. In: *Digital Scholarship in the Humanities*, 31 (1), S. 118–139.
- Krippendorff, Klaus (1980): *Content analysis: An introduction to its methodology*. Beverly Hills u.a.: Sage.
- Krumm, Hans-Jürgen/Fandrych, Christian/Hufeisen, Britta/Riemer, Claudia (Hrg.) (2010): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*. Berlin u.a.: de Gruyter. (= Handbücher zur Sprach- und Kommunikationswissenschaft).
- Krummes, Cedric/Ensslin, Astrid (2012): Formulaic language and collocations in German essays. From corpus-driven data to corpus-based materials. In: *The Language Learning Journal* 43 (1), S. 110-127.
- Kubczak, Jaqueline (2009): Hier wird Ihnen geholfen! E-VALBU - Das elektronische Valenzwörterbuch deutscher Verben. In: *Sprachreport* 4, S. 17-23.
- Kuhberg, Heinz (1990): Zum L2-Erwerb zweier elfjähriger Kinder mit Türkisch und Polnisch als Ausgangssprachen. In: *Deutsch lernen* 15, S. 25-43.
- Kwakernaak, Erik (2005): Kasusmarkierungen bei niederländischsprachigen Deutschlernenden. Entwurf eines Erwerbsszenarios. In: *Deutsch als Fremdsprache* 4, S. 222-231.
- Kuhberg, Heinz (2001): Zweitspracherwerb als prädestinierte Entwicklung I: der behavioristische Ansatz. In: Helbig et al (Hrg), S. 654-663.
- Lado, Robert (1957): *Linguistics across Cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Larsen-Freeman, Diane (2009): Another step to be taken: Rethinking the endpoint of the interlanguage continuum. In: Han, ZhaoHong/Tarone, Elaine (Hrg.): *Interlanguage: Forty Years Later*. Amsterdam: John Benjamins. S. 203–220.
- Larson-Hall, Jenifer (2010): *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York/London: Routledge.
- LCB = Centre for English Corpus Linguistics (2019): *Learner Corpus Bibliography*. Louvain-la-Neuve: Université catholique de Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpus-bibliography.html> (Mai 2019).
- LCW = Centre for English Corpus Linguistics (2019): *Learner corpora around the world*. Louvain-la-Neuve: Université catholique de Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html> (Mai 2019).
- LdN (2019): *Lexikon der Neurowissenschaft*. Heidelberg: Spektrum Akademischer Verlag. <https://www.spektrum.de/lexikon/neurowissenschaft/> (Mai 2019)
- Leacock, Claudia/Chodorow, Martin/Gamon, Michael/Tetreault, Joel (2014): *Automated Grammatical Error Detection for Language Learners*. Second edition. Morgan & Claypool Publishers. (= Synthesis Lecture on Human Language Technologies 9).
- Leacock, Claudia/Chodorow, Martin/Tetreault, Joel (2015): Automatic grammar- and spell-checking for language learners. In: Granger/Gilquin/ Meunier (Hrg.), S. 567-586.
- Lefcheck, Jonathan S. (2016): piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. In: *Methods in Ecology and Evolution* 7 (5), S. 573-579.
- Lennon, Paul (1991): Error. Some Problems of Definition, Identification, and Discussion. In: *Applied Linguistics* 12 (2), S. 180-196.

- Lemnitzer, Lothar/Zinsmeister, Heike (2015): Korpuslinguistik. Eine Einführung. 3. Auflage. Tübingen: Narr.
- Lerot, Jaques (1982): Die verbregierten Präpositionen in Präpositionalobjekten. In: Abraham, Werner (Hrg.): Satzglieder im Deutschen. Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung. Tübingen: Narr. S. 261-291.
- Leys, Odo (1989): Aspekt und Rektion räumlicher Präpositionen. In: Deutsche Sprache 2, S. 97-113.
- Lightbrown, Patsy M./Spada, Nina (2013): How Languages are Learned. Oxford: Oxford UP.
- Little, David (2006): The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. In: Language Teaching 39 (3). S. 167-190.
- Lüdeling, Anke (2008): Mehrdeutigkeiten und Kategorisierung: Probleme bei der Annotation von Lernerkorpora. In: Walter, Maik/Grommes, Patrick (Hrg.): Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitspracherwerbsforschung. Tübingen: Niemeyer. S. 119-140.
- Lüdeling, Anke/Doolittle, Seanna/Hirschmann, Hagen/Schmidt, Karin/Walter, Maik (2008): Das Lernerkorpus Falko. In: Deutsch als Fremdsprache 2, S. 67-73.
- Lüdeling, Anke/Hirschmann, Hagen (2015): Error annotation systems. In: Granger/Gilquin/Meunier (Hrg.), S. 135-158.
- Lüdeling, Anke/Hirschmann, Hagen/Shadrova, Anna (2017): Linguistic Models, Acquisition Theories, and Learner Corpora: Morphological Productivity in SLA Research Exemplified by Complex Verbs in German. In: Language Learning 67, S. 1-34.
- Lüdeling, Anke/Kytö, Merja (Hrg.) (2008): Corpus Linguistics. An International Handbook. Band 1. Berlin u.a.: Mouton de Gruyter.
- Lüdeling, Anke/Walter, Maik (2009): Korpuslinguistik für Deutsch als Fremdsprache. Sprachvermittlung und Spracherwerbsforschung. https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/anke/anke_veroeffentlichungen (Mai 2019).
- Lüdeling, Anke/Walter, Maik (2010): Korpuslinguistik. In: Krumm et al. (Hrg.), S. 315-322.
- Lüdeling, Anke/Walter, Maik/Kroymann, Emil/Adolphs, Peter (2005): Multi-level error annotation in learner corpora. In: Proceedings of Corpus Linguistics 2005, 14.-17. Juni. Birmingham. <http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-e-journal.aspx> (Mai 2019).
- Lütke, Beate (2008): Beobachtung zur Raumreferenz in mündlichen Erzählungen. In: Ahrenholz, Bernt (Hrg.): Zweitspracherwerb. Diagnosen, Verläufe, Voraussetzungen. Freiburg: Fillibach. S. 151-170.
- Lütke, Beate (2010): Lokale Präpositionalphrasen in syntaktischer Perspektive – ein hypothetisches Progressionsmodell auf der Basis von Lernerdaten. In: Rost-Roth, Martina (Hrg.): DaZ-Spracherwerb und Sprachförderung Deutsch als Zweitsprache. Beiträge aus dem 5. Workshop "Kinder mit Migrationshintergrund". Freiburg u.a.: Fillibach. S. 33-52.
- Lütke, Beate (2011): Deutsch als Zweitsprache in der Grundschule: Eine Untersuchung zum Erlernen lokaler Präpositionen. Berlin u.a.: de Gruyter.
- MacWhinney, Brian (1987): The Competition Model. In: MacWhinney, Brian (Hrg.): Mechanisms of Language Acquisition. Hillsdale: Lawrence Erlbaum. S. 249-308.
- MacWhinney, Brian (2001): The competition model: the input, the context and the brain. In: Robinson, Peter (Hrg.): Cognition and Second Language Instruction. Cambridge: Cambridge UP, S. 69-90.

- Maden-Weinberger, Ursula (2008): Modality as Indicator of L2 Proficiency? A corpus-based investigation into advanced German interlanguage. In: Walter, Maik/Grommes, Patrick (Hrg.): Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitspracherwerbsforschung. Tübingen: Niemeyer. S. 141-164. (= Linguistische Arbeiten 50).
- Maden-Weinberger, Ursula (2015): "Hätte, wäre, wenn...". A pseudo-longitudinal study of subjunctives in the Corpus of Learner German (CLEG). In: International Journal of Learner Corpus Research 1, S. 25-57.
- Madhani, Nitin/Tetreault, Joel/Chodorow, Martin/Rozovskaya, Alla (2011): They Can Help: Using Crowdsourcing to Improve the Evaluation of Grammatical Error Detection Systems. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics: Portland, Oregon, Juni 2011. S. 508-513.
- Marossek, Diana (2016): Kommst du Bahnhof oder hast du Auto? Warum wir reden, wie wir neuerdings reden. München: Hanser Berlin Verlag.
- Mater, Erich (1971): Deutsche Verben. Bd. 6. Leipzig: VEB Bibliographisches Institut.
- Meindl, Claudia (2011): Methodik für Linguisten. Eine Einführung in Statistik und Versuchsplanung. Tübingen: Narr.
- Meisel, Jürgen (1977): Linguistic simplification: a study of immigrant workers' speech and foreigner talk. In: Corder, Stephen P./Roulet, Eddy (Hrg.): The notion of simplification, interlanguages and pidgins and their relation to second language pedagogy. Actes du 5. Colloque de Linguistique Appliquée de Neuchâtel, 20-22 Mai 1976. Genève. S. 99-113.
- MERLIN project (2014): Annotation guidelines. www.merlin-platform.eu (Mai 2019).
- MERLIN Richtlinien (2014) = MERLIN project, Richtlinien für die Digitalisierung/Transkription der Lernerdaten im MERLIN-Projekt (2014). <http://merlin-platform.eu> (Mai 2019).
- Meunier, Fanny (2015): Developmental Patterns in Learner Corpora. In: Granger/Gilquin/Meunier (Hrg.), S. 379-400.
- Meurers, Detmar (2015): Learner Corpora and Natural Language Processing. In: Granger/Gilquin/Meunier (Hrg.), S. 537-566.
- Meyer-Ingwersen, Johannes (1975): Einige typische Deutschfehler bei türkischen Schülern. In: Zeitschrift für Literaturwissenschaft und Linguistik 5, S. 68-77.
- Minitab Blog (2013): Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit? <http://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit> (Mai 2019).
- Mitchell, Rosamond/Myles, Florence/Marsden, Emma (2013): Second Language Learning Theories. Third Edition. Abingdon: Routledge.
- Müller, Christoph/Strube, Michael (2006): Multi-Level Annotation of Linguistic Data with MMAX2. In: Braun, Sabine/Kohn, Kurt/Mukherjee, Joybrato (Hrg.): Corpus Technology and Language Pedagogy. New Resources, New Tools, New Methods. Frankfurt: Peter Lang. S. 197-214. (= English Corpus Linguistics 3).
- Myles, Florence (2015): Second language acquisition theory and learner corpus research. In: Granger/Gilquin/Meunier (Hrg.), S. 309-332.
- Nesselhauf, Nadja (2004): Learner corpora and their potential for language teaching. In: Sinclair, John (Hrg.): How to Use Corpora in Language Teaching. Amsterdam/Philadelphia: John Benjamins. S. 125-152.
- Ott, Margarete (2006): Wortschatzerwerb und Erwerb grammatischer Strukturen. Eine empirische Studie zum Zweitspracherwerb Deutsch. In: Deutsch als Fremdsprache 3,

- S. 163-171.
- Palmberg, Rolf (1980): A select bibliography of error analysis and interlanguage studies. Abo: Abo Akademi.
- Pepouna, Soulemanou/Nsangou, Maryse/Ndassi, Laurent/Nono, Arnaud (2018): Der muttersprachliche Hintergrund als Einflussfaktor auf den Objektkasuserwerb in der Fremdsprache Deutsch. In: Deutsch als Fremdsprache 2, S. 71-81.
- Pfaff, Carol W. (1984): On Input and Residual L1 Transfer Effects in Turkish and Greek Children's German. In: Andersen, Roger (Hrg.): Second Languages. A cross-linguistic perspective. Rowley, MA: Newbury House. S. 271-298.
- Philippi, Jule (2008): Einführung in die generative Grammatik. Göttingen: Vandenhoeck & Ruprecht.
- Pica, Teresa (1984): Methods of Morpheme Quantification: Their Effect on the Interpretation of Second Language Data. In: Studies of Second Language Acquisition 6 (1), S. 69-78.
- Pienemann, Manfred (1981): Der Zweitspracherwerb ausländischer Arbeiterkinder. Bonn: Bouvier.
- Pienemann, Manfred (1984): Psychological Constraints on the Teachability of Languages. In: Studies in Second Language Acquisition 6 (2), S. 186-214.
- Pienemann, Manfred (1989): Is Language teachable? Psycholinguistic Experiments and Hypotheses. In: Applied Linguistics 10 (1), S. 52-79.
- PONS (2019): Feste Verbindungen: Verben mit Präpositionen. Stuttgart: PONS GmbH. <https://de.pons.com/shop/specials/zusatzinhalte-produkte/praxisgrammatik/deutsch> (Dezember 2019).
- Primus, Beatrice (1999): Rektionsprinzipien. In: Wegener, Heide (Hrg.): Deutsch kontrastiv: typologisch-vergleichende Untersuchungen zur deutschen Grammatik. Tübingen u.a.: Stauffenburg. S. 135-170.
- Primus, Beatrice (2012): Semantische Rollen. Heidelberg: Winter.
- Proost, Kristel (2015): Verbbedeutung, Konstruktionsbedeutung oder beides? Zur Bedeutung deutscher Ditransitivstrukturen und ihrer präpositionalen Varianten. In: Engelberg, Stefan/Meliss, Meike/Proost, Kristel/Winkler, Edeltraud (Hrg.): Argumentstruktur zwischen Valenz und Konstruktion. Tübingen: Narr. S. 157-176. (=Studien zur deutschen Sprache 68)
- R Core Team (2018): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (Mai 2019).
- Rankin, Tom (2015): Learner Corpora and Grammar. In: Granger/Gilquin/Meunier (Hrg.), S. 231-254.
- Ragheb, Marwa/Dickinson, Markus (2013): Inter-annotator Agreement for Dependency Annotation of Learner Language. In: Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications. Juni 2013. Atlanta, Georgia: Association for Computational Linguistics. S. 169-179. <https://aclweb.org/anthology/volumes/proceedings-of-the-eighth-workshop-on-innovative-use-of-nlp-for-building-educational-applications/> (Mai 2019).
- Rastelli, Stefano (2009): Learner Corpora without Error Tagging. In: Linguistic Online 38 (2), S. 57-66. http://www.linguistik-online.com/38_09/rastelli.html (Mai 2019).
- Rauh, Gisa (1999): Language-specific mental models: The case of English *on* and German *auf*. In: Missler, Bettina/Multhaou, Uwe (Hrg.): The Construction of Knowledge, Learner Autonomy and Related Issues in Foreign Language Learning (Essays in Honour of Dieter Wolff). Tübingen: Stauffenburg. S. 109-125.
- Reznicek, Marc (2012): Falko-Excel-AddIn. <https://www.linguistik.hu->

- berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/marc/standardseite (Mai 2019).
- Reznicek, Marc/Lüdeling, Anke/Hirschmann, Hagen (2013): Competing Target Hypotheses in the Falko Corpus. A Flexible Multi-Layer Corpus Architecture. In: Díaz-Negrillo, Ana/Ballier, Nicolas/Thompson, Paul (Hrg.): Automatic Treatment and Analysis of Learner Corpus Data. Amsterdam u.a.: John Benjamins. S. 101-123.
- Reznicek, Marc/Lüdeling, Anke/Krummes, Cedric/Schwantuschke, Franziska/Walter, Maik/Schmidt, Karin/Hirschmann, Hagen/Andreas, Torsten (2012): Das Falko-Handbuch: Korpusaufbau und Annotationen: Version 2.01. Berlin. <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2/view> (Mai 2019).
- Reznicek, Marc/Zinsmeister, Heike (2013): STTS-Konfusionsklassen beim Tagging von Fremdsprachlernertexten. In: Journal for Language Technology and Computational Linguistics 28 (1), S. 63-83.
- Richard Lowry (2001-2009): VassarStats: Website for Statistical Computation. <http://vassarstats.net/index.html> (Mai 2019).
- Rindler Schjerve, Rosita (1981): Zu den Präpositionen DI, A, DA in einer kontrastiven Grammatik des Italienischen und des Deutschen. In: Italienische Studien 4, S. 127-131.
- Roe, Ian F. (2007): Valency and the errors of learners of English and German. In: Herbst, Thomas/Götz-Votteler, Katrin (Hrg.) (2007): Valency. Theoretical, Descriptive and Cognitive Issues. Berlin u.a.: de Gruyter. S. 217-228.
- Rostila, Jouni (2014): Inventarisierung als Grammatikalisierung. Das Beispiel Präpositionalobjekte. In: Lasch, Alexander/Ziem, Alexander (Hrg.): Grammatik als Netzwerk von Konstruktionen. Sprachwissen im Fokus der Konstruktionsgrammatik. Berlin/Boston: de Gruyter. S. 97-116. (= Sprache und Wissen (SuW) 15).
- Rostila, Jouni (2018): Argument structure constructions among German prepositional objects. In: Boas, Hans C./Ziem, Alexander (Hrg.): Constructional Approaches to Syntactic Structures in German. Berlin/Boston: De Gruyter. S. 406-446. (= Trends in Linguistics. Studies and Monographs (TiLSM) 322).
- Rundell, Michael/Granger, Sylviane (2007): From corpora to confidence. In: English Teaching Professional 50, S. 15-18.
- Rösch, Heidi (2011): Deutsch als Zweit- und Fremdsprache. Berlin: Akademie Verlag.
- RStudio, Inc. (2018): Open source and enterprise-ready professional software for R. <https://www.rstudio.com/> (Mai 2019).
- Schiller, Anne/Teufel, Simone/Stöckert, Christine (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). Universität Tübingen. <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/GermanTagsets.html> (Mai 2019).
- Schmid, Helmut (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees. In: Proceedings of International Conference on New Methods in Language Processing, Manchester, UK.
- Schmid, Helmut (2008): Tokenizing and part-of-speech tagging. In: Lüdeling/Kytö (Hrg.), S. 527-551.
- Schmidt, Claudia (2008): Grammatik und Korpuslinguistik. Überlegungen zur Unterrichtspraxis DaF. In: Deutsch als Fremdsprache 2, S. 74-80.
- Schmidt, Claudia (2010): Kognitivistische/Konstruktivistische/Konnektionistische Ansätze. In: Krumm et al. (Hrg.), S. 807-817.

- Schmidt, Karin (2010): Lernerkorpora: Ressourcen für die Deutsch-als-Fremdsprache-Forschung. In: Eđit, Yadigar (Hrg.): Globalisierte Germanistik: Sprache, Literatur, Kultur: Tagungsbeiträge. XI. Türkischer Internationaler Germanistik-Kongress 20.- 22. Mai 2009. Izmir: Ege Üniver. Matbaasi, S. 555-573.
- Schmitt, Richard/Dreyer, Hilke (2000): Lehr- und Übungsbuch der deutschen Grammatik. Ismaning: Hueber.
- Schroeder, Christoph/Şişmek, Yazgül (2010): Kontrastive Analyse Türkisch-Deutsch. In: Krumm et al. (Hrg.), S. 719-725.
- Schwarze, Christoph (2011): Grammatik der italienischen Sprache. 2., verbesserte Auflage. Berlin u.a.: de Gruyter.
- Selinker, Larry (1972): Interlanguage. In: International Review of Applied Linguistics in Language Teaching (IRAL) 10 (3), S. 209-231.
- Sieradz, Magdalena/Bordag, Denisa (2014): Zum Gebrauch von Verschmelzungen aus Präposition und bestimmten Artikel bei Deutschlernenden und Muttersprachlern. Eine empirische Untersuchung. In: Deutsch als Fremdsprache 51, S. 11.18.
- Skiba, Romuald (2008): Korpora in der Zweitspracherwerbsforschung. Internetzugang zu Daten des ungesteuerten Zweitspracherwerbs. In: Ahrenholz, Bernt/Bredel, Ursula/Klein, Wolfgang/Rost-Roth, Martina/Skiba, Romuald (Hrg.): Empirische Forschung und Theoriebildung. Beiträge aus der Soziolinguistik, Gesprochene-Sprache- und Zweitspracherwerbsforschung. Berlin: Lang. S. 21-30.
- Steinmüller, Ulrich (1992): Spracherwerbsbiographie und Zweisprachigkeit. In: Wissenschaftliche Zeitschrift der Humboldt-Universität Berlin. Reihe Geistes- und Sozialwissenschaften 5, S. 25.40. <http://www.daf.tu-berlin.de/fileadmin/fg75/Artikel/SprachBio.pdf> (Mai 2019).
- Storrer, Angelika (1992): Verbvalenz. Theoretische und methodische Grundlagen ihrer Beschreibung in Grammatikographie und Lexikographie, Tübingen: Niemeyer (= Reihe Germanistische Linguistik 126).
- Storrer, Angelika (1996): Wie notwendig sind obligatorische Valenzstellen? - Zur Weglaßbarkeit von Valenzstellen im Text. In: Gréciano, Gertrud/Schumacher, Helmut (Hrg.): Lucien Tesnière - Syntaxe structurale et opérations mentales. Akten des deutsch-französischen Kolloquiums anläßlich der 100. Wiederkehr seines Geburtstages. Straßburg 1993. Niemeyer: Tübingen. S. 225-238.
- Storrer, Angelika (2003): Ergänzungen und Angaben. In: Ágel et al. (Hrg.), S. 764-780.
- Sylla, Bernard (1999): Zum Problem der Kasuswahl nach Wechselpräposition. In: Deutsch als Fremdsprache 3, S. 150-155.
- Tagliamonte, Sali/Baayen, R. Harald (2012): Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. In: Language Variation and Change 24 (2), S. 135-178.
- Tarone, Elaine (2013): Interlanguage. In: Capelle, Carol (Hrg.): The Encyclopedia of Applied Linguistics 6, Ho-La. Chichester: Wiley-Blackwell. S. 2794-2801.
- Telljohann, Heike/Hinrichs, Erhard W./Kübler, Sandra/Zinsmeister, Heike/Beck, Kathrin (2015): *Stylebook for the Tübingen Treebank of Written German (TüBa-D/Z)*. Revised Version. Technical Report, Seminar für Sprachwissenschaft, Universität Tübingen. <http://www.sfs.uni-tuebingen.de/ascl/ressourcen/corpora/tueba-dz.html> (Mai 2019).
- Tenfjord, Kari/Hagen, Jon E./Johansen, Hilde (2006): The hows and whys of coding categories in a learner corpus (or "how and why an error tagged learner corpus is not IPSO FACTO one big comparative fallacy"). In: Rivista Di Psicolinguistica Applicata 6 (3), S. 93-108.

- Tetreault, Joel/Chodorow, Martin (2008a): Native Judgements of non-native usage: Experiments in preposition error detection. In: Coling 2008: Proceedings of the workshop on Human Judgements in Computational Linguistics. Manchester, August 2008. S. 24-32. <http://www.aclweb.org/anthology/W08-12> (Mai 2019).
- Tetreault, Joel/Chodorow, Martin (2008b): The Ups and Downs of Preposition Error Detection in ESL Writing. In: Coling 2008: Proceedings of the 22nd International Conference on Computational Linguistics. Manchester, August 2008. S. 865-872. <https://www.aclweb.org/anthology/C08-1000> (Mai 2019).
- Tetreault, Joel/Filatova, Elena/Chodorow, Martin (2010): Rethinking grammatical error annotation and evaluation with the Amazon Mechanical Turk. In: Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications, Los Angeles, California, June 2010. Association for Computational Linguistics, S. 45-48. <https://www.aclweb.org/anthology/papers/W/W10/W10-1006/> (Mai 2019).
- Tippmann, Sylvia (2014): *Programming tools: Adventures with R*. A guide to the popular, free statistics and visualization software that gives scientists control of their own data analysis. In: Nature 517 (1), S. 109-110. <https://www.nature.com/news/programming-tools-adventures-with-r-1.16609> (Mai 2019).
- Tschirner, Erwin (2006): Häufigkeitsverteilungen im Deutschen und ihr Einfluss auf den Erwerb des Deutschen als Fremdsprache. In: Corino, Elisa/Marello, Carla/Onesti, Cristina (Hrg.): Atti del XII Congresso Internazionale di Lessicografia. Alessandria: Edizioni dell'Orso, S. 1277-1288.
- Turgay, Katharina (2010a): Der Zweitspracherwerb der deutschen Präpositionalphrase: eine Studie zum Sprachentwicklungsstand von Kindern mit Migrationshintergrund. Trier: WVT. (= Koblenz-Landauer Studien zu Geistes-, Kultur- und Bildungswissenschaften 8).
- Turgay, Katharina (2010b): Unterspezifikation von Präpositionen - eine Studie zum Erwerb von PPs mit Wechselpräpositionen. In: Pohl, Inge (Hrg.): Semantische Unbestimmtheit im Lexikon. Frankfurt a.M. u.a.: Peter Lang. S. 223-243. (= Sprache. System und Tätigkeit 61).
- Turgay, Katharina (2010c): Lokale Adpositionalkonstruktionen im Deutschen und Türkischen Ein formaler und funktionaler Vergleich. In: Deutsche Sprache 4, S. 309-325.
- Turgay, Katharina (2011): Der Erwerb des deutschen Kasus in der Präpositionalphrase. In: Zeitschrift für Germanistische Linguistik 3, S. 24-54.
- Uhlisch, Gerda (1995): Immer wieder: Interferenz(fehler) im Fremdsprachenunterricht. In: Handwerker, Brigitte (Hrg.): Fremde Sprache Deutsch: grammatische Beschreibung - Erwerbsverläufe - Lehrmethodik. Tübingen: Narr. S. 225-238.
- Universität Leipzig (2017): SpraStu. Sprache und Studienerfolg bei Bildungsausländer/-innen. Eine empirische Längsschnittstudie an den Universitäten Leipzig und Würzburg. <http://home.uni-leipzig.de/sprastu/projekt/#disko> (Mai 2019).
- VanPatten, Bill/Williams, Jessica (2015) (Hrg.): Theories in second language acquisition: An Introduction. Second Edition. New York u.a.: Routledge.
- van Rooy, Bertus (2015): Annotating learner corpora. In: Granger/Gilquin/Meunier (Hrg.), S. 79-106.
- Viorel, Elena (1973): Schwierigkeiten der deutsch lernenden Rumänen beim Gebrauch von Präpositionen. In: Deutsch als Fremdsprache 6, S. 354-357.
- Vyatkina, Nina (2016): The Kansas Developmental Learner Corpus (KANDEL). A developmental corpus of learner German. In: International Journal of Learner Corpus Research 2 (1), S. 102-120.

- Walter, Maik/Doolittle, Seanna (2012): Vorfeldfortschritte im Deutschen als Fremdsprache. Eine lernerkorpusbasierte Longitudinalstudie amerikanischer Deutschlerner. In: *Deutsch als Fremdsprache* 1, S. 3-11.
- Walter, Maik/Grommes, Patrick (2008): Die Entdeckung des fortgeschrittenen Lerners in der Varietätenlinguistik. In: Walter, Maik/Grommes, Patrick (Hrg.): *Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitspracherwerbsforschung*. Tübingen: Niemeyer. S. 3-28. (= *Linguistische Arbeiten* 50).
- Weber, Tassja (2014): *Verbvalenz und Rektion im Bereich Deutsch als Fremdsprache. Eine korpusgestützte Analyse zweier Verbgruppen* (Masterarbeit TU Dortmund). https://merlin-platform.eu/C_download.php (Mai 2019).
- Weber, Tassja (2015): *Verb Valency and Prepositional Complements in Learner Corpora: A Case Study in the German MERLIN Corpus*. In: de Haan, Pieter (Hrg.): *LCR 2015 Book of Abstracts*. S. 164-166.
- Weber, Tassja (2018): *Grammatik und Lernerkorpora: Eine korpusorientierte Untersuchung von Präpositionalphrasen im deutschen MERLIN-Korpus*. In: Fuß, Eric/Konopka, Marek/Trawiński, Beata/Waßner, Ulrich H. (Hrg.): *Grammar and Corpora 2016*. Heidelberg: Heidelberg University Publishing. S. 415-424. DOI: <https://doi.org/10.17885/heiup.361.509>.
- Weber, Tassja (2019): *A Corpus-based Approach to the Usage and Acquisition of Prepositions by Learners of German as a Foreign Language: Form vs. function*. In: Abel, Andrea /Glaznieks, Aivars/Lyding, Verena/Nicolas, Lionel (Hrsg.): *Widening the Scope of Learner Corpus Research. Selected Papers from the Fourth Learner Corpus Research Conference*. Louvain-la-Neuve: Presses universitaires de Louvain. S. 121-136. (= *Corpora and Language in Use*)
- Wegener, Heide (1995a): *Kasus und Valenz im natürlichen DaZ-Erwerb*. In: Eichinger, Ludwig/Eroms, Hans Werner (Hrg.): *Dependenz und Valenz*. Hamburg: Buske. S. 337-356.
- Wegener, Heide (1995b): *Das Genus im DaZ-Erwerb. Beobachtungen an Kindern aus Polen, Rußland und der Türkei*. In: Handwerker, Brigitte (Hrg.): *Fremde Sprache Deutsch: grammatische Beschreibung – Erwerbsverläufe – Lehrmethodik*. Tübingen: Narr. S. 1-24. (= *Tübinger Beiträge zur Linguistik* 409).
- Wegener, Heide (1995c): *Die Nominalflexion des Deutschen – verstanden als Lerngegenstand*. Tübingen: Niemeyer. (=Reihe *Germanistische Linguistik* 151).
- Wegener, Heide (2008): *Der Erwerb eines komplexen morphologischen Systems in DaZ: Der Plural deutscher Substantive*. In: Walter/Grommes (Hrg.), S. 93-118.
- Wiese, Heike/Pohle, Maria (2016): „Ich geh Kino“ oder „... ins Kino“? *Gebrauchsrestriktionen nicht kanonischer Lokalangaben*. In: *Zeitschrift für Sprachwissenschaft* 35 (2), S. 171-216.
- Willems, Klaas (2006): *Ist Valenz eine universelle Eigenschaft von Verben?* In: *Zeitschrift für Dialektologie und Linguistik* 73(1), S. 55-76.
- Wilmots, Jos/Moonen, Erik (1997): *Der Gebrauch von Akkusativ und Dativ nach Wechselpräpositionen*. In: *Deutsch als Fremdsprache* 3, S. 144-149.
- Wisniewski, Katrin (2017): *Das Potenzial von Lernerkorpora im Sprachtestbereich*. In: *Deutsch als Fremdsprache* 1, S. 33-40.
- Wisniewski, Katrin (2018a): *Studienerfolg und sprachliche Kompetenzen bei Bildungsausländer/-innen: Das Forschungsprojekt SpraStu*. In: *Info Deutsch als Fremdsprache* 45 (4), S. 573–597 (= *Themenheft Schlüssel zum Bildungserfolg: Sprachliche Kompetenzen fördern und bewerten*).

- Wisniewski, Katrin (2018b): Grammatische Korrektheit und L2-Kompetenz: Eine Lernerkorpus-Studie. In: *Deutsch als Fremdsprache* 3, S. 131-142.
- Wisniewski, Katrin/Woldt, Claudia/Schöne, Karin/Abel, Andrea/Blaschitz, Verena/ Štindlová, Barbara/Vodičková, Kateřina (2014): The MERLIN annotation scheme for the annotation of German, Italian, and Czech learner language. www.merlin-platform.eu (Mai 2019).
- Wolfer, Sascha (2017): *Verstehen und Verständlichkeit juristisch-fachsprachlicher Texte* Tübingen: Narr. (=Korpuslinguistik und interdisziplinäre Perspektiven auf Sprache, CLIP 7)
- Wolfer, Sascha/Hansen-Morath, Sandra (2018): Visualisierung sprachlicher Daten mit R. In: Bubenhofer, Noah/Kupietz, Marc (Hrg.): *Visualisierung sprachlicher Daten. Visual Linguistics – Praxis – Tools*. Heidelberg: Heidelberg University Publishing. S. 227-255. DOI: <https://doi.org/10.17885/heiup.345.474> (Mai 2019).
- Wray, Alison (2002): *Formulaic Language and the Lexicon*. Cambridge: CUP.
- Zeldes, Amir/Lüdeling, Anke/Hirschmann, Hagen (2008): What's hard? Quantitative evidence for difficult constructions in German learner data. In: *Proceedings of QITL 3*. Helsinki. DOI: <https://doi.org/10.18452/13445> (Mai 2019).
- Zeldes, Amir/Ritz, Julia/Lüdeling, Anke/Chiarcos, Christian (2009): ANNIS: a search tool for multi-layer annotated corpora. In: Mahlberg, Michaela/González-Díaz, Victorina/Smith Catherine (Hrg.): *Proceedings of the Corpus Linguistics Conference (CL)*, July 20-23, Liverpool. <http://ucrel.lancs.ac.uk/publications/cl2009/> (Mai 2019).
- Zinsmeister, Heike/Breckle, Margit (2012): The ALeSKo learner corpus. Design – annotation – quantitative analyses. In: Schmidt, Thomas/Wörner, Kai (Hrg.): *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam/Philadelphia: Benjamins. S. 71-96.
- Zinsmeister, Heike/Reznicek, Marc/Ricart Brede, Julia/Rosén, Christina/Skiba, Dirk (2012): Das Wissenschaftliche Netzwerk "Kobalt-DaF". In: *Zeitschrift für Germanistische Linguistik* 40 (3), S. 457-458.
- Zipser, Florian/Romary, Laurent (2010): A model oriented approach to the mapping of annotation formats using standards. In: *Proceedings of the Workshop on Language Resource and Language Technology Standards, LREC 2010*. Malta. <http://hal.archives-ouvertes.fr/inria-00527799/en/> (Mai 2019).

Quellenverzeichnis

MERLIN: Das Lernerkorpus MERLIN. Universität Tübingen/EURAC Bozen. <https://merlin-platform.eu/>.

Anhang

Im Anhang der Dissertation finden sich folgende Unterlagen:

1. Anhang 1: Outputs der statistischen Auswertungen (Fallstudie 1, 2 und 3)
2. Anhang 2: Annotationshandbuch (Fallstudie 2 und 3)

Anhang 1: Outputs der statistischen Auswertungen

Fallstudie 1

RM1

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Korrektheit

	Chisq	Df	Pr(>Chisq)
L1.Zf	2.3168	1	0.1280
Kompetenzniveau	25.6464	2	2.697e-06
L1.Zf:Kompetenzniveau	1.9714	2	0.3732

RM1b

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Korrektheit

	Chisq	Df	Pr(>Chisq)
Kompetenzniveau	24.802	2	4.115e-06

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Korrektheit	binomial	logit	delta	0.01941311	0.04263937

RM2a

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Wahl

	Chisq	Df	Pr(>Chisq)
L1.Zf	10.7277	1	0.001055
Kompetenzniveau	5.6556	2	0.059143 .

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Wahl	binomial	logit	delta	0.135238	0.135238

RM3a

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Realisierung	binomial	logit	delta	0.09205347	0.09205347

RM4

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Wahl	binomial	logit	delta	0.007875933	0.007875933

RM5

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Wahl	binomial	logit	delta	0	0

Fallstudie 2

RM1

Output Regressionsmodell

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.3712	0.2410	5.691	1.26e-08
PP_FunktionO	-2.1522	0.4067	-5.292	1.21e-07
Kompetenzniveau.opB1	0.2059	0.3064	0.672	0.501637
Kompetenzniveau.opB2	0.2674	0.3068	0.872	0.383401
Kompetenzniveau.opC1	1.3364	0.8141	1.642	0.100650
PP_FunktionO:Kompetenzniveau.opB1	1.5283	0.4937	3.096	0.001962
PP_FunktionO:Kompetenzniveau.opB2	1.8135	0.4988	3.636	0.000277
PP_FunktionO:Kompetenzniveau.opC1	1.7120	1.1051	1.549	0.121328

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Korrektheit

	Chisq	Df	Pr(>Chisq)
PP_Funktion	18.889	1	1.385e-05
Kompetenzniveau.op	18.548	3	0.000339
PP_Funktion:Kompetenzniveau.op	14.241	3	0.002595

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Korrektheit	binomial	logit	delta	0.06679877	0.1690051

RM2

Output Anova()

Response: Auslassung

	Chisq	Df	Pr(>Chisq)
PP_Funktion	15.7485	1	7.234e-05
Kompetenzniveau.op	19.9531	2	4.648e-05
PP_Funktion:Kompetenzniveau.op	1.4926	2	0.4741

RM3

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Wahl

	Chisq	Df	Pr(>Chisq)
PP_Funktion	12.7876	1	0.0003489
Kompetenzniveau.op	16.1171	2	0.0003164
PP_Funktion:Kompetenzniveau.op	0.3427	2	0.8425233

RM2a

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Auslassung

	Chisq	Df	Pr(>Chisq)
PP_Funktion	16.216	1	5.653e-05
Kompetenzniveau.op	20.523	2	3.496e-05

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Auslassung	binomial	logit	delta	0.2629708	0.314703

RM3a

Output Anova()

Response: Wahl

	Chisq	Df	Pr(>Chisq)
PP_Funktion	13.020	1	0.0003083
Kompetenzniveau.op	16.402	2	0.0002744

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Wahl	binomial	logit	delta	0.1715339	0.2157364

Fallstudie 3

RM1

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Korrektheit

	Chisq	Df	Pr(>Chisq)
Praeposition	13.2478	1	0.0002729
Kompetenzniveau.op	40.1631	3	9.84e-09
Praeposition:Kompetenzniveau.op	2.7605	3	0.4300444

RM1a

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Korrektheit

	Chisq	Df	Pr(>Chisq)
Praeposition	13.763	1	0.0002074
Kompetenzniveau.op	42.618	3	2.966e-09

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Korrektheit	binomial	logit	delta	0.1177381	0.1758991

RM2

Output Anova()

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: Hinzufuegung

	Chisq	Df	Pr(>Chisq)
Praeposition	42.7464	1	6.232e-11
Kompetenzniveau.op	0.2814	2	0.8687
Praeposition:Kompetenzniveau.op	4.4479	2	0.1082

RM2b

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Hinzufuegung	binomial	logit	delta	0.2755648	0.2755648

RM3a

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Auslassung	binomial	logit	delta	0.1843438	0.1843438

RM4a

Output rsquared()

Response	family	link	method	Marginal	Conditional
1 Wahl	binomial	logit	delta	0.1350049	0.135004

Anhang 2: Annotationshandbuch

Anmerkung:

Das vorliegende Annotationshandbuch wurde für die manuelle Annotationen der PP-Kontexte in Fallstudie 2 und 3 erstellt. Die für Annotationszwecke eingesetzte Version des Handbuchs wurde für die Einreichung der Dissertation minimal überarbeitet. Zu den Überarbeitungen zählen vor allem die Sicherstellung einer einheitlichen Terminologie und Formatierung sowie die Neu-Anordnung einiger (Unter-)Kapitel. Die zentralen Inhalte des Annotationshandbuchs (vom Stand Juni 2018) wurden nicht verändert.

Das Annotationshandbuch wurde primär für den Gebrauch durch zwei studentische Hilfskräfte am Lehrstuhl Germanistische Linguistik der Universität Mannheim verfasst. Bei der Entwicklung des Annotationshandbuchs konnte ich ein grundlegendes linguistisches Wissen seitens der Annotatorinnen voraussetzen. Das Annotationshandbuch dient primär der Darstellung des Annotationsprozesses und der Erklärung einzelner Annotationskategorien für die Sicherstellung einer einheitlichen und konsistenten Annotation durch die beiden Hilfskräfte.

Die Inhalte des Annotationshandbuchs wurden stetig, parallel zu den Annotationsphasen und Diskussionen, überarbeitet und ggf. durch weitere Informationen oder Beispiele, etc. ergänzt. In seiner vorliegenden Form haben sich die Inhalte des Annotationshandbuchs für die Annotationen der Fallstudie 2 und 3 bewährt.

Annotationshandbuch

Inhaltsverzeichnis

1. Vorwort: Hinweise zum Annotationsdokument	292
2. Annotationskategorien	293
2.1 Syntaktische Funktionen der PP	293
2.2 Erläuterungen zu den syntaktischen Funktionen.....	293
2.2.1 Objekt-PP.....	293
2.2.1.1 Ressourcen zum Nachschlagen: E-Valbu und Duden-Online.....	295
2.2.2 Adverbiale PP.....	301
2.2.3 Attributive PP.....	302
2.2.4 Annotation der syntaktischen Funktionen von PP: Hilfestellung	304
2.2.5 Annotation der syntaktischen Funktionen von PP: Ergänzende Hinweise für PP des Fehlertyps <i>Hinzufügung</i>	304
2.3 Weitere Annotationskategorien.....	306
2.3.1 Rest (R)	306
2.3.2 Pseudotreffer.....	307
2.3.3 Nicht interpretierbar (NI).....	307
2.3.4 Funktionsverbgefüge (FVG)	308
2.3.5 Formelhafte Sequenz (FS).....	309
2.3.6 Kommentar	309
2.4 Annotationskategorien und –werte im Überblick.....	309
2.4.1. Ausgewählte Annotationsbeispiele.....	312

2.5 Annotation Fehlertyp im Bereich Präpositionsgebrauch	313
2.5.1 Annotation Fehlertyp: Überblick.....	314
2.5.2 Allgemeine Hinweise zur Annotation des Fehlertyps im Präpositionsgebrauch.....	315
2.5.3 Annotation Fehlertyp: Umgang mit fehlenden Fehlerannotationen.....	317
2.5.4 Annotation Fehlertyp: Umgang mit inkorrekten Fehlerannotationen	317
2.6 Annotation Gewähltes Verb/Adjektiv	318
3. Ergänzende Hinweise zur Nachrecherche im Lernerkorpus MERLIN.....	319
4. Annotationen: Abschließende Hinweise und Checkliste	320
5. Literaturverzeichnis und Ressourcen	322
Anhang zum Annotationshandbuch.....	323

1. Vorwort: Hinweise zum Annotationsdokument

Die Annotationen erfolgen in mehreren (Excel-)Annotationsdokumenten. Pro Präposition und GeR-Gesamtniveau liegt ein Annotationsdokument (= ein Excel-Dokument) vor. Das Annotationsdokument darf auf keinen Fall verändert werden, Zeilen bzw. Spalten o.ä. dürfen nicht (!) geändert oder gelöscht werden, weder inhaltlich noch formal. Die oberen Zeilen in jedem Dokument dienen der automatischen Auswertung. Bei der Annotation können diese ausgeblendet werden. Jedes Annotationsdokument enthält die jeweils exportierten PP-Kontexte aus dem Lernerkorpus MERLIN. Die Zeile *tok* repräsentiert die Original-Lerneräußerung, die Zeile *TH1*³⁰¹ ist die minimale Zielhypothese der entsprechenden Original-Lerneräußerung, d. h. sie stellt eine Rekonstruktion der Original-Lerneräußerung dar, wobei zielsprachliche Abweichungen in den Bereichen Grammatik und Orthographie korrigiert wurden. An dieser Stelle sei bereits erwähnt, dass die Annotationswerte jeweils auf der Ebene der TH1 (=selbe Zeilenhöhe wie die TH1) eingetragen werden. Beim Eintrag der Annotationswerte bitte darauf achten, dass hinter den einzelnen Annotationswerten kein Leerzeichen ist. Dies ist für die automatische Auszählung der Annotationswerte sehr wichtig.

Weiterhin ist zu beachten, dass die Annotationsdokumente einen Namen haben, der nicht beliebig geändert werden darf. Wenn annotierte Annotationsdokumente ausgetauscht werden, bitte folgendes im Dateinamen der jeweiligen Datei ergänzen: (...)_annotiert_Namenskürzel_Datum.xlsx. Dabei dient als Namens Kürzel der jeweils erste Buchstabe des Vornamens und des Nachnamens. Beispiel: (...)_annotiert_TW_10.01.17.xlsx. Wichtig: Bitte Unterstriche (_) und keine Leerzeichen im Dateinamen benutzen.

³⁰¹ *TH1* ist die englische Abkürzung für die minimale Zielhypothese (ZH1). Im Annotationshandbuch wird mit *TH1* auf die ZH1 verwiesen, da die jeweiligen Annotationsdokumente ebenfalls die Abkürzung TH1 enthalten.

2. Annotationskategorien

2.1 Syntaktische Funktionen der PP³⁰²

Zu annotieren ist vorrangig die **syntaktische Funktion der Präpositionalphrase (PP)**, wie sie in der Original-Lerneräußerung bzw. auf TH1-Ebene gebraucht wird; dabei berücksichtigt man die Original-Lerneräußerung (Ebene *tok*) und gleicht diese mit der Ebene der Zielhypothese (*TH1*) ab. In manchen Fällen fehlt auf tok-Ebene eine Präposition (und damit die entsprechende PP), wobei diese auf Ebene der TH1 vorhanden ist, da sie in der Zielhypothese ergänzt wurde. In diesen Fällen wird die syntaktische Funktion der PP entsprechend ihrer Verwendungsweise in der ZH1 annotiert³⁰³. Zu den Fällen, in denen die Präposition bzw. PP auf tok-Ebene vorhanden ist, auf TH1-Ebene jedoch gelöscht wurde, siehe Kap. 2.2.5.

Als PP gelten Strukturen, die aus einer Präposition und einem Nominal bestehen. Dabei kann das Nominal eine Nominalphrase oder z. B. eine Pronominalphrase sein. Beispiele: [*mit dir*], [*für dich*], [*auf dem schön gedeckten Tisch*], [*an der Haustür*], etc.

Je nach syntaktischem Kontext können die PP in unterschiedlichen syntaktischen Funktionen (z. B. *Objekt*, *Adverbiale*, *Attribut*) gebraucht werden. Diese syntaktischen Funktionen sollen annotiert werden.

Die syntaktischen Funktionen, die hier angesetzt werden, sind:

1. Objekt-PP (hierzu gehören auch PP, die von Adjektiven regiert werden. Diese werden jedoch in einer separaten Spalte annotiert)
2. Adverbiale PP (PP in der Funktion eines Adverbiale)
3. Attributive PP (PP in der Funktion eines Attributs)
(vgl. auch Duden 2016: 851ff.)

2.2 Erläuterungen zu den syntaktischen Funktionen

2.2.1 Objekt-PP

Unter einer Objekt-PP versteht man eine PP in der Funktion eines Objekts. Die PP ist abhängig von einem Verb und die Präposition wird vom Verb festgelegt, d. h. die Präposition wird vom Verb regiert. Es liegt somit Präpositionsreaktion vor.

Beispiele für Verben mit Objekt-PP³⁰⁴:

³⁰² Die Beispiele für den PP-Gebrauch wurden u.a. entnommen aus GDS (1997), Eisenberg (2013) und Duden (2016), z. T. stammen Beispiele auch von der Autorin.

³⁰³ Dabei darf die TH1 nicht verändert werden.

³⁰⁴ Das entsprechende Verb ist jeweils in Klammern angegeben. Die Lesarten der jeweiligen Verben sind hier natürlich ebenfalls zu beachten; so fordert das Verb *leben* zum Beispiel in einer ganz bestimmten Lesart die PP *für etw./jmdn.* (Lesart: *sich jemandem/etwas widmen*), in einer anderen Lesart (*irgendwo seinen Wohnsitz haben*) dahingegen fordert das Verb diese PP nicht. Aus diesem Grund ist eine Recherche in E-Valbu (näher Informationen und Hinweise s. unten) sehr hilfreich, wenn es darum geht, zu bestimmen, ob eine Objekt-PP vorliegt.

- *Ich warte auf meine Mutter* (warten auf)
- *Ich kümmerge mich um den Abwasch* ((sich) kümmern um)
- *Ich passe auf die Hunde auf* (aufpassen auf)
- *Ich stimme für Partei X* (stimmen für)
- *Ich interessiere mich für Partei X* ((sich) interessieren für)
- *Ich spreche mit meiner Schwester* (sprechen mit)
- *Ich rechne mit einer Gehaltserhöhung* (rechnen mit)
- *Das Symbol steht für Liebe* (stehen für)
- *Ich halte seinen Vorschlag für sehr vernünftig* (halten für)
- *Sie lebt nur für ihre Kinder* (leben für)
- *Wir verstehen uns mit den Nachbarn* ((sich) verstehen mit)
- *Er brach mit seinen Eltern* (brechen mit)
- *Ich danke dir für deinen Brief* (danken für)
- *Er kommuniziert mit seinen Mitarbeitern* (kommunizieren mit³⁰⁵)

Bei Objekt-PP handelt es sich um PP, die in der Bedeutung des Verbs ‚vorangelegt‘ sind (vgl. Duden 2016: 779). Beispielsweise impliziert das Verb *warten*, dass es etwas oder jemanden gibt, auf den gewartet wird. Die Handlung ist auf etwas oder jemanden gerichtet: *Ich warte auf meine Eltern*. Auch wenn man die PP *auf meine Eltern* weglassen würde, würde sie dennoch in der Bedeutung des Verbs implizit mitverstanden werden.

Im Gegensatz dazu sind adverbiale PP i.d.R. nicht in der Bedeutung des Verbs angelegt. Das Verb *warten* impliziert z. B. nicht, wie lange oder wo die Handlung des Wartens vollzogen wird. Diese Informationen können jedoch ergänzt werden, z. B. *Ich warte am Bismarckplatz auf meine Eltern*. Die PP *am Bismarckplatz* ist dann eine adverbiale PP. Näheres zu adverbial gebrauchten PP s. weiter unten.

Die Präposition in der Objekt-PP ist in der Regel³⁰⁶ nicht austauschbar und wenn man die PP als Ganzes ersetzen möchte, ist dies i.d.R. nur durch ein Pronominaladverb möglich, das sich aus *wo(r)/da* entsprechender P zusammensetzt, oder durch eine andere präpositionsartige Fügung (vgl. Duden 2016: 853, Eisenberg 2013: 302). Beispiele:

- *Ich warte auf meine Mutter -> ich warte auf Sie*
- *Ich kümmerge mich um den Abwasch -> ich kümmerge mich darum*
- *Ich passe auf die Hunde auf -> ich passe auf Sie auf*
- *Ich passe auf die Schuhe auf -> ich passe darauf auf*
- *Ich stimme für mehr Gleichberechtigung -> ich stimme dafür*
- *Ich rechne mit einer Gehaltserhöhung -> ich rechne damit*

³⁰⁵ In Anlehnung an *sprechen (mit)* wird die *mit*-PP beim Verb *kommunizieren* als Objekt-PP klassifiziert.

³⁰⁶ Dies gilt tatsächlich nicht für alle Objekt-PP, da z. B. beim Verb *stimmen für* die Präposition ausgetauscht werden kann (*stimmen gegen*). Jedoch kommen bei diesem Verb nur zwei Präpositionen infrage, andere sind nicht möglich.

Wichtig 1:

Zu den Objekt-PP werden bei der Annotation auch PP gefasst, die die Funktion von Prädikativen einnehmen (prädikative PP). Charakteristisch für prädikative PP ist, dass die PP in prädikativer Funktion „eine Aussage über eine oder mehrere andere Phrasen“ (Duden 2016: 799) macht. Diese andere Phrase, über die eine Aussage getroffen wird, wird Bezugsphrase genannt. Diese kann entweder das Subjekt oder das Objekt des Satzes sein, beim Prädikativ kann also Subjekt- oder Objektbezug vorliegen.

Beispiel für eine prädikative PP mit Objektbezug:

Sie erklärte den Patienten für gesund

-> (hier bezieht sich die PP auf das Akkusativobjekt *Patienten*, die PP wird vom Verb regiert und die Präposition ist nicht austauschbar)

Wichtig 2:

PP können nicht nur von Verben regiert werden, sondern auch von Adjektiven. D. h. es gibt Adjektive (s. u.), die eine bestimmte Präposition regieren (Duden 2016: 851f.). PP, die von Adjektiven regiert werden, werden bei der Annotation separat unter der Kategorie *Objekt-PP Adjektiv (OAdj)* annotiert. Diese PP werden bei der Auswertung zu der Kategorie Objekt-PP gezählt.

In der Regel stehen diese Adjektive in Verbindung mit einem Kopulaverb.

Beispiele:

- *stolz (sein) auf*
- *entsetzt (sein) über*
- *geübt (sein) in*
- *interessiert (sein) an*
- *gedacht (sein) für*
- *geeignet (sein) für*
- *notwendig (sein) für*

Hinweis:

Bei PP-Kontexten, die als *Objekt-PP* bzw. *Objekt-PP Adjektiv* annotiert werden, soll anschließend auch das gewählte Verb bzw. Adjektiv annotiert werden (s. Kap. 2.6).

2.2.1.1 Ressourcen zum Nachschlagen: E-Valbu und Duden-Online

Zur Kontrolle, ob es sich bei den PP um Objekt-PP handelt, bitte die elektronische Ressource des Valenzwörterbuchs deutscher Verben, *E-Valbu*³⁰⁷, konsultieren³⁰⁸. E-Valbu sollte bei der

³⁰⁷ <http://hypermedia.ids-mannheim.de/evalbu/index.html> (bis November 2018).

³⁰⁸ Die Screenshots im Annotationshandbuch sowie die entsprechenden Hinweise zur Suche in E-Valbu stammen von der ursprünglichen Website von E-Valbu. Seit Dezember 2018 ist E-Valbu

Annotation noch nicht dokumentierter Verben³⁰⁹ immer konsultiert werden. E-Valbu enthält Informationen zu Verblesarten und Satzbauplänen³¹⁰, wobei der Fokus auf der valenzgrammatischen Perspektive liegt. D. h. in E-Valbu findet man Informationen darüber, ob ein Verb eine PP-Ergänzung regiert. Ist diese PP-Ergänzung eine Objekt-PP, so lautet die entsprechende Abkürzung in E-Valbu *Kprp*. Da in der Annotation prädikative PP ebenfalls zu Objekt-PP gezählt werden und dementsprechend als Objekt-PP annotiert werden, ist auch die E-Valbu-Abkürzung für prädikative PP relevant. Die entsprechende Abkürzung in E-Valbu lautet *Kprd*³¹¹. Weiterhin sind adverbiale PP – aber nur solche, die gleichzeitig PP-Ergänzungen sind – mit der Abkürzung *Kadv* benannt.

Im Folgenden wird näher auf die Recherche in E-Valbu eingegangen.

Hinweise zur Recherche in E-Valbu:

Auf der Startseite von E-Valbu (Abb. 1) kann oben im Reiter der Anfangsbuchstabe des Verbs gesucht werden (einfach auf Buchstabe draufklicken), die Liste mit entsprechenden Verben erscheint dann links. Dort kann man durch einen Klick das gewünschte Verb auswählen.



Abbildung 1: Startseite E-Valbu (Screenshot).

Sobald das entsprechende Verb ausgewählt wurde, gelangt man zum Verbartikel. Dort sind u.a. die unterschiedlichen Lesarten des Verbs aufgeführt. Die Lesarten sind nummeriert. Der folgende Screenshot (Abb. 2) zeigt ein entsprechendes Beispiel für das Verb *achten*:

ausschließlich über [grammis \(https://grammis.ids-mannheim.de/verbvalenz\)](https://grammis.ids-mannheim.de/verbvalenz) erreichbar. Die Inhalte haben sich, so wurde mir nach einer Rückfrage bestätigt, jedoch nicht verändert.

³⁰⁹ Eine Dokumentation von Verben, die Objekt-PP fordern, findet sich in diesem Handbuch (Kap. 2.2.1) und auch im Dokument *Beispiele Annotationen* (s. Anhang). Die entsprechenden Beispiele sind während der Annotation zu berücksichtigen.

³¹⁰ Der Satzbauplan gibt Informationen darüber, welche Komplemente ein Verb fordert und welche Form diese Komplemente einnehmen können (vgl. Duden 2016: 927ff.).

³¹¹ Näheres zu den Begrifflichkeiten in E-Valbu unter <https://grammis.ids-mannheim.de/systematische-grammatik/2912>.

achten	
Aussprache:	'achten
Stammformen:	achtet - achtete - hat geachtet
Konjugationsmuster:	schwach
1 achten auf	seine Aufmerksamkeit auf die Einhaltung von etwas richten
2 achten	jemanden/etwas schätzen
3 achten	etwas respektieren
4 achten auf	auf etwas aufpassen
5 achten auf	etwas bewusst wahrnehmen
6 achten auf	etwas in seine Überlegungen einbeziehen

Abbildung 2: E-Valbu: Verb "achten" (Screenshot).

In Abbildung 2 sieht man, dass für einige Lesarten das Verb mit einer Präposition aufgeführt ist und für andere nicht. Das bedeutet, dass je nach Lesart das Verb eine obligatorische Objekt-PP fordert oder eben nicht. Für die Bestimmung der syntaktischen Funktion einer PP ist folglich zunächst die Bestimmung der Lesart eines Verbs im zu annotierenden Datenausschnitt sehr wichtig. Dabei ist zu beachten, dass obligatorische Ergänzungen bzw. PP direkt in der Auflistung der unterschiedlichen Verblesarten aufgeführt werden, so wie oben im Beispiel von *achten auf* (Lesart 1/4/5/6). Fakultative Ergänzungen (also solche, die nicht obligatorisch sind), sind in der Übersicht der Verblesarten nicht aufgeführt, s. folgendes Beispiel für das Verb (*sich*) *bedanken* (Lesart 1) (s. Abb. 3).

1 sich [A] bedanken	jemandem gegenüber mittels irgendetwas seinen Dank für etwas ausdrücken
---------------------	---

Abbildung 3: E-Valbu: Verb "(sich) bedanken" (Screenshot).

Hier (Abb. 3) ist die PP (mit der Präposition *für*) nicht in der Übersicht zur Verblesart aufgeführt, jedoch ist sie in der Paraphrase der Bedeutung des Verbs enthalten (... *für etwas*). Im vollständigen Artikel (vgl. Abb. 4) findet man einen Verweis darauf, dass die *für*-PP eine fakultative Ergänzung des Verbs *sich bedanken* darstellt (s. Pfeile in Abb. 4). Fakultative Ergänzungen sind im Satzbauplan in Klammern angegeben.


sich [A] bedanken 1 (lesartspezifische Angaben) [Artikelk]	
Strukturbeispiel:	jemand bedankt sich bei jemandem mit irgendetwas für etwas ←
Im Sinne von:	jemand drückt jemandem gegenüber mittels irgendetwas seinen Dank für etwas aus
Satzbauplan:	K _{sub} , (K _{prp}) , (K _{prp}) , (K _{adv})
Beispiele:	(1)  Mit einem großen Blumenstrauß und einer Urkunde bedankte sich Schwechats Polizeidirektor Franz Schnabl im Rahmen einer kleinen Feier bei Anna Graf für ihre mutige Tat. (Neue Kronen-Zeitung, 12.01.1999, S. 18)

Abbildung 4: E-Valbu: Verb "(sich) bedanken" (Lesart 1), Auszug aus vollständigem Artikel (Screenshot).

In E-Valbu erhält man folglich vollständigen Informationen zur Lesart und zum Satzbauplan des Verbs, wenn man auf die jeweilige Lesarten-Nr. klickt. Dies sollte immer gemacht werden; ein Blick auf die Übersicht der unterschiedlichen Verblesarten genügt in der Regel nicht, um eindeutig zu bestimmen, ob die Lesart, die in E-Valbu angegeben ist, auch mit der Lesart übereinstimmt, die man bei einem bestimmten PP-Kontext vorliegen hat.

Hier noch ein Beispiel aus E-Valbu für das Verb *achten* in der Lesart *auf jemanden aufpassen* (Lesart 4) (Abb. 5). Unter *Satzbauplan* findet man die Kategorien der Ergänzungen, die das Verb in der entsprechenden Lesart registriert. Abbildung 5 zeigt, dass das Verb *achten* in der Lesart *auf jemanden aufpassen* (u. a.) eine Objekt-PP (*Kprp*) fordert.

achten auf 4 (lesartspezifische Angabe)	
Strukturbeispiel:	[geh] jemand achtet auf jemanden/etwas
Im Sinne von:	jemand passt auf jemanden/etwas auf; Acht geben
Satzbauplan:	K _{sub} , K _{prp}

Abbildung 5: E-Valbu: Verb "achten auf" (Lesart 4), Auszug aus vollständigem Artikel (Screenshot).

Im Folgenden wird auf prädikative PP eingegangen, die bei der Annotation ebenfalls zu den Objekt-PP gezählt werden. Die Kategorie *Kprd* steht für die Kategorie *Prädikativergänzung*. Diese Ergänzung kann u. a. in der Form einer PP repräsentiert sein, z. B. im Satz: *Ich halte ihn für inkompetent*. Die folgende Abbildung (Abb. 6) gibt ein Beispiel aus E-Valbu für das Verb *halten für*, das eine prädikative PP fordert.

halten für 1 (lesartspezifische Angaben) [Artike	
Strukturbeispiel:	jemand hält jemanden/etwas für jemanden/etwas
Im Sinne von:	jemand schätzt jemanden/etwas als jemanden/etwas ein oder ist der Meinung, dass jemand/etwas so ist
Satzbauplan:	K _{sub} , K _{akk} , K _{prd}
Beispiele:	(1) Halten Sie die Menschen für lernfähig? (Zeit, 30.08.1985, S. 36) (2) Die Regierung hält die Steuerreform für realisierbar.

Abbildung 6: E-Valbu: Verb "halten für" (Lesart 1), Auszug aus vollständigem Artikel (Screenshot).

Wie oben bereits erwähnt, werden PP in Objekt- als auch in Prädikativfunktion unter der Kategorie *Objekt-PP* annotiert.

E-Valbu kann auch genutzt werden, um adverbiale PP zu identifizieren. Wie oben erwähnt, werden in E-Valbu Ergänzungen zu Verben (bzw. Verblesarten) angegeben. Handelt es sich bei dieser Ergänzung um ein Adverbiale, so ist dieses mit der Abkürzung *Kadv* benannt. In Abbildung 4 findet sich ein Beispiel für *Kadv*. Das Verb (*sich*) *bedanken für* regiert eine Ergänzung in der Funktion eines Adverbiale. In Abbildung 4 findet man dafür das Beispiel *Mit einem großen Blumenstrauß*. Hier liegt folglich eine adverbiale *mit-PP* vor. Im Annotationsdokument sind folglich ähnliche PP-Kontexte als adverbiale PP zu annotieren. Zur Erinnerung: In E-Valbu findet man lediglich Informationen zu adverbialen PP, die gleichzeitig Ergänzungen von Verben sind. Adverbiale PP, die keine Ergänzungen sind, können auf diese Weise nicht identifiziert werden. Für die Identifikation von adverbialen PP, die keine Ergänzungen sind, ist es sinnvoll, die Informationen unter Kap. 2.2.2 zu beachten bzw. das Dokument *Beispiele Annotationen* zu nutzen (s. Anhang).

Wichtig:

Bitte im Annotationsdokument in der Spalte „Kommentar“ vermerken, wenn E-Valbu als Ressource konsultiert wurde („E-Valbu, Lesart X³¹²“).

Ist ein Verb nicht in E-Valbu enthalten, wird das entsprechende Verb in einem Universal-Wörterbuch nachgeschlagen, z. B. im Duden Universalwörterbuch bei Duden-Online³¹³. Dort können in der Regel die Umschreibung der Bedeutung des Verbs und/oder die Verwendungsbeispiele des Verbs Hinweise darauf geben, ob das Verb eine Präposition bzw. eine Objekt-PP regiert.

Beispiel: Das Verb *zurechtkommen* (s. Abb. 7). In der Umschreibung der Verbbedeutung wird zunächst deutlich, dass das Verb ein Objekt fordert. Die Umschreibung der Bedeutung des Verbs und vor allem die Verwendungsbeispiele enthalten jeweils eine *mit-*

³¹² An dieser Stelle dann bitte auch die entsprechende Lesart-Nr. einsetzen.

³¹³ <http://www.duden.de/woerterbuch>.

PP. Es kann daher angenommen werden, dass das Verb *zurechtkommen* eine Objekt-PP (mit der Präposition *mit*) regiert.

1. für etwas ohne große Schwierigkeiten einen möglichen Weg, die richtige Lösung finden, es bewältigen; mit jemandem, etwas fertigwerden

Beispiele

- wie soll man mit einer solchen Maschine zurechtkommen?
- sie kommen mit den Kindern nicht mehr zurecht

Abbildung 7: Duden-Online: Verb "zurechtkommen" (Bedeutung 1). Bedeutung und Verwendungsbeispiele (Screenshot).

Wichtig:

Im Annotationsdokument in der Spalte „Kommentar“ vermerken, wenn Duden als Ressource konsultiert wurde („Duden, Bedeutung X³¹⁴“).

Kategorie Objekt-PP Adjektiv

Hinweise zur PP, die von Adjektiven regiert werden, findet man z. B. in Duden (2016: 366).

Beispiele:

Die Lehrerin ist mit den Leistungen zufrieden -> zufrieden (sein) mit

Diese Behörde ist in Bonn ansässig -> ansässig (sein) in

Zudem findet man Hinweise zur PP, die von Adjektiven regiert werden, auch in einem Universal-Wörterbuch, z. B. Duden Universalwörterbuch oder Duden-Online (s.o.). Dort liefert in der Regel die Umschreibung der Bedeutung des Adjektivs bzw. die Verwendungsbeispiele des Adjektivs einen Hinweis darauf, ob das Adjektiv eine Präposition bzw. eine Objekt-PP regiert.

Beispiel: Das Adjektiv *interessiert (sein)* (Abb. 8). Die Verwendungsbeispiele des Adjektivs enthalten zwei Beispiele mit *an*-PP und ein Beispiel mit einem Pronominaladverb (*daran*). Die Verwendungsbeispiele zeigen, dass das Adjektiv *interessiert* (im nicht-attributivem Gebrauch innerhalb einer NP) eine *an*-PP regiert (s. Pfeile in Abb. 8).

³¹⁴ An dieser Stelle dann bitte auch die entsprechende Nr. einsetzen.

[starken] Anteil nehmend; Interesse habend, zeigend; geistig aufgeschlossen, aufmerksam

Beispiele

- ein interessierter junger Mann
- ein interessantes Gesicht machen
- sie ist literarisch, politisch interessiert
- interessiert zuhören
- ➔ • an diesem Mädchen sind viele interessiert (*haben viele Interesse*)
- ➔ • sie ist an diesem Problem nicht interessiert (*es interessiert sie nicht*)
- ➔ • wir sind sehr daran interessiert, dieses Geschäft zu machen (*wir möchten es sehr gerne machen*)

Abbildung 8: Duden-Online: Adjektiv "interessiert". Bedeutung und Verwendungsbeispiele (Screenshot).

2.2.2 Adverbiale PP

Adverbiale PP dienen, grob gesagt, dazu, einen Sachverhalt zu situieren (Eisenberg 2013: 183). Als ein Adverbiale ist eine PP bzw. die Präposition innerhalb der PP in der Regel nicht vom Verb regiert. Die PP ist daher oft weglassbar (vgl. Duden 2016: 852f.).

Beispiele:

- *Ich esse einen Apfel in der Küche.*
- *Ich esse einen Apfel hinter dem Haus.*

Es gibt jedoch auch Adverbialien (bzw. adverbiale PP), die vom Verb regiert sind und aus diesem Grund nicht weglassbar sind (vgl. Duden 2016: 793 /Eisenberg 2013: 303f.), Beispiel: *Sie begibt sich auf die Bühne.*

Bei der Annotation wird nicht (!) zwischen weglassbaren und nicht weglassbaren adverbialen PP unterschieden. Beide sind jeweils als „adverbiale PP“ zu annotieren.

Semantisch betrachtet können adverbiale PP verschiedene Funktionen übernehmen:

1) Adverbiale PP können sich „von außen“ entweder auf eine Äußerung als Ganzes beziehen und z. B. die Sprechereinstellungen zu einem Sachverhalt ausdrücken (auch Kommentaradverbialien oder Satzadverbialien genannt, s. Duden 2016: 794):

Für mein Gefühl ist das falsch.

Anna kommt mit Sicherheit zu spät.

2) Adverbiale PP können sich „von innen“ auf das Prädikat beziehen oder auf den Sachverhalt als Ganzes (auch situative Adverbialien genannt, Duden 2016: 795):

Sie schnitt die Wurst mit einem Messer.

Sie schnitt die Wurst in der Küche.

Bei adverbialen PP ist die Präposition in der PP i.d.R. austauschbar und wenn man die PP als Ganzes ersetzen möchte, ist dies i.d.R. durch ein einfaches Adverb möglich. Die entsprechende

Präposition muss, anders als bei Objekt-PP, nicht ebenfalls realisiert werden (vgl. Duden 2016: 853):

- *Ich wohne in Mannheim -> Ich wohne hier.*
- *Ich esse einen Apfel in der Küche -> Ich essen einen Apfel dort.*
- *ich esse einen Apfel an diesem schönen Tag. -> Ich esse einen Apfel dann/heute.*

Es gibt jedoch auch Ausnahmen, vgl. *Ich aß den Fisch mit einem Messer -> Ich aß ihn damit*
In solchen Beispielen wie dem obigen (*mit dem Messer*) ist die PP ebenfalls durch das Pronominaladverb *damit* ersetzbar, das ist also identisch zur Objekt-PP mit *mit* (*Ich rechne damit*). Im Gegensatz zu *rechnen mit* ist die PP bei *essen* aber nicht vom Verb regiert bzw. führt das Weglassen der PP *mit dem Messer* nicht zu einem Unterschied in der Bedeutung des Verbs. Bei *rechnen* ergibt sich allerdings eine andere Verbbedeutung, wenn die PP weggelassen wird, vgl. *Ich rechne mit einer Gehaltserhöhung* vs. *Ich rechne*. Im Beispielsatz *Ich aß den Fisch mit einem Messer* wäre zudem (wenn auch nicht elegant klingend) ein Ersatz mit einem Adverb möglich (vgl. *Ich esse so*)³¹⁵.

Beispiele für adverbiale PP sind:

- *Ich esse einen Apfel auf der Parkbank.*
- *Ich esse einen Apfel am Rhein.*
- *Ich arbeite mit meinem besten Freund.*
- *Ich arbeite ohne meinen besten Freund.*
- *Ich arbeite neben dem Chefbüro.*
- *Wir prüfen das mit Sorgfalt*
- *Er fährt für zwei Jahre nach Amerika*
- *Ich wünsche dir viel Glück für die Prüfung*
- *Ich habe eine Fahrkarte für einen Wochenendausflug gekauft.*

Weitere Beispiele finden sich im Dokument *Beispiele Annotationen* (s. Anhang). Dieses Dokument bitte während der Annotationen berücksichtigen.

2.2.3 Attributive PP

Attributive PP sind PP in der Funktion eines Attributs. Attribute modifizieren in der Regel Substantive (s. u.).

Wichtig: Bei der Annotation von attributiven PP werden nur PP innerhalb von Nominalphrasen (NP) und Pronominalphrasen als attributive PP klassifiziert (vgl. auch Duden 2016: 790, 852).
Beispiele für attributive PP innerhalb einer NP sind:

- *Das Haus auf dem Berg*
- *Die Antwort auf meine Frage*
- *Hilfe für die Kinder*

³¹⁵ Zur Ersetzbarkeit von PP durch Adverbien s. auch Eisenberg (2013: 302).

- Einfluss auf die Gesellschaft

Als Attribut ist die PP Bestandteil eines Satzglieds, d. h., dass eine attributive PP kein eigenständiges Satzglied ist, nicht frei verschiebbar ist und folglich (in der Regel) auch in einem Verbzweitsatz nie alleine vor dem finiten Verb stehen kann.

Beispiel: Die NP *Der Mann mit dem Hut* beinhaltet die attributive PP *mit dem Hut*:

[Der Mann [mit dem Hut]] schaute zu mir herüber.

**[Mit dem Hut] schaute [der Mann] zu mir herüber.*

**[Der Mann] schaute [mit dem Hut] zu mir herüber.*

Weitere Beispiele:

- *Das Haus [auf dem Berg] ist marode.*
- *Eine Hoffnung [auf gutes Wetter] hatte sie nicht mehr*

In den obigen Sätzen ist die attributive PP ebenfalls durch die Position erkennbar: Es steht gemeinsam mit dem Satzglied, von dem es abhängt, vor dem finiten Verb. Da in der Regel nur ein Satzglied vor dem finiten Verb stehen kann, muss es sich bei *[auf dem Berg]* im Satz *Das Haus [auf dem Berg] ist marode* folglich um eine eingebettete PP und damit um eine attributive PP handeln.

Im Gegensatz dazu kann man im Satz *Mit einem Lächeln begrüßte mich die Frau* erkennen, dass es sich bei der PP *Mit einem Lächeln* um keine attributive PP handelt, da die PP als eigenständiges Satzglied alleine vor dem finiten Verb steht.

TIPP:

Ein sinnvoller Test zur Ermittlung einer attributiven PP ist die Verschiebung der PP vor das finite Verb (in einem Verbzweitsatz). Verändert sich dadurch die Aussage des Satzes bzw. wird der Satz ungrammatisch, so handelt es sich bei der PP in der Regel um eine attributive PP (vgl. auch Duden 2016: 789f.)

An dieser Stelle sei jedoch darauf hingewiesen, dass es hier auch Ausnahmen gibt: Im folgenden Satz liegt ebenfalls eine attributive PP vor, aber ein Verschieben vor das finite Verb ist möglich: *Sie hatte einen großen Einfluss [auf die Gesellschaft]. -> [Auf die Gesellschaft] hatte sie einen großen Einfluss.*

Man spricht im Fall derartiger PP auch von der Integration der PP ins Prädikat und sagt dann, dass die PP von der gesamten Konstruktion *einen Einfluss haben* abhängt (vgl. Duden 2016: 790f.). Dennoch handelt es sich hier bei der PP auf die Gesellschaft um eine attributive PP (vgl. Duden 2016: 868), da die PP das Nomen (*Einfluss*) näher spezifiziert (und die PP ebenfalls weggelassen werden kann): Es geht hier um einen bestimmten Einfluss, und zwar den auf die Gesellschaft.

PP in Sätzen wie z. B. *Sie hatte einen großen Einfluss [auf die Gesellschaft]* werden daher als attributive PP annotiert.

Ergänzender Hinweis für attributive *am*-PP:

PP, die auf eine mehrfache Wiederholung einer Tätigkeit im Laufe des Tages, der Woche o.ä. Bezug nehmen, werden als attributive PP annotiert.

Beispiel: *füttere den zwei Mal am Tag* -> die *am*-PP bezieht sich auf die Häufigkeit, mit der das Füttern ausgeführt werden soll und spezifiziert diese näher.

Analog werden folgende *am*-PPs als attributive PP annotiert:

x-mal am Tag

x Stunden am Tag

spät am Abend

2.2.4 Annotation der syntaktischen Funktionen von PP: Hilfestellung

Bei der Annotation der einzelnen syntaktischen Funktionen der PP können folgende Schritte hilfreich sein:

- 1) Bestimmung des (finiten) Verbs im vorliegenden, zu annotierendem Ausschnitt bzw. Satz.
- 2) Kontrolle des Verbs und der Lesart des Verbs in E-Valbu. Wenn Verb nicht in E-Valbu vorliegt, Kontrolle in einem Universalwörterbuch (Duden-Online).
- 3) Wenn keine Objekt-PP vorliegt: Bestimmung, ob eine attributive oder eine adverbiale PP vorliegt: Kontrolle der Position der PP (allein-/nicht alleinstehend vor dem finiten Verb?, Anwendung des Verschiebe-Tests, s. o.) und Interpretation der Bedeutung der PP (wird durch die PP etwas näher bestimmt? Wird durch die PP eine Satzaussage kommentiert? Wird sich durch die PP auf die Verbhandlung bezogen?).

2.2.5 Annotation der syntaktischen Funktionen von PP: Ergänzende Hinweise für PP des Fehlertyps *Hinzufügung*

In den Annotationsdokumenten finden sich z. T. einige PP-Kontexte, in denen eine Präposition in der Original-Lerneräußerung (tok) enthalten ist, auf der Ebene der TH1 jedoch gelöscht wurde. Diese Kontexte repräsentieren eine hinzugefügte Präposition, die zielsprachlich jedoch nicht gefordert ist. Die PP-Kontexte, die diesen Fehlertyp („Hinzufügung“) repräsentieren, finden sich, wenn vorliegend, ganz unten in jedem Annotationsdokument. Auch diese PP-Kontexte werden gemäß den Annotationskategorien, die in diesem Kapitel vorgestellt werden, annotiert. Für die Annotation der syntaktischen Funktionen von PP solcher Fälle muss Folgendes beachtet werden:

- 1) Für die Annotation der syntaktischen Funktion der entsprechenden PP ist ausschlaggebend, wie die jeweilige PP gebraucht wurde und welche Phrase sich an entsprechender Stelle auf Ebene der TH1 findet.
- 2) Für die Annotation der syntaktischen Funktion der entsprechenden PP bitte an der TH1-Ebene orientieren. Es ist also zu beachten, an welcher Stelle die Präposition auf der tok-Ebene vorkommt und welche Phrase sich an genau dieser Stelle auf Ebene der TH1 findet. Die syntaktische Funktion der Einheit, die der PP auf tok-Ebene entspricht, wird interpretiert und die syntaktische Funktion dieser Einheit wird dann annotiert. Die folgenden Beispiele illustrieren das Vorgehen bei der Annotation:

1.

Tok: *Ich habe [an Silvester] mit X verbracht.*

TH1: *Ich habe [Silvester] mit X verbracht.*

-> In diesem Beispiel liegt eine Präposition (*an*) vor, die auf Ebene der TH1 gelöscht wurde. Die Stelle *an Silvester* wurde auf TH1 zu *Silvester* geändert. Im Satz auf der Ebene der TH1 *Ich habe [Silvester] mit X verbracht* stellt die NP *Silvester* das Objekt im Satz dar (Akkusativobjekt zum Verb *verbringen*). Man kann die Modifikation von tok zu TH1 also dergestalt interpretieren, dass die auf tok produzierte *an*-PP anstelle eines Kasusobjekts realisiert wurde. Das bedeutet für die Annotation der syntaktischen Funktion dieses PP-Kontexts, dass die PP die syntaktische Funktion eines Objekts ausdrückt. Dieser PP-Kontext ist folglich als Objekt-PP zu annotieren.

2.

Tok: *bitt schreibst [für mich].*

TH1: *Bitte schreibe [mir].*

-> In diesem Beispiel liegt eine Präposition (*für*) vor, die auf Ebene der TH1 gelöscht wurde. Die PP *für mich* wurde auf TH1 zu *mir* geändert. D. h. die Konstruktion *für mich* auf tok-Ebene entspricht der Konstruktion *mir* auf TH1-Ebene. Die Einheit *mir* erfüllt im Satz *Bitte schreibe mir* die syntaktische Funktion des Objekts. Folglich wird dieser PP-Kontext als Objekt-PP annotiert.

Achtung Kategorie „Nicht interpretierbar“ (NI):

3.

Tok: *Ein Balkon auf .*

TH1: *, einen Balkon _____.*

In *Beispiel 3* liegt ebenfalls eine hinzugefügte Präposition vor, auf der Ebene der TH1 findet man jedoch an dieser Stelle keine Entsprechung. D. h. es gibt hier auf TH1-Ebene keine Phrase, die man für Annotationszwecke interpretieren kann. Solche Fälle werden als Kategorie „nicht interpretierbar“ annotiert.

Ähnliches gilt für PP-Kontexte, in denen neben der hinzugefügten Präposition auf tok-Ebene kein Verb enthalten ist, auf TH1 jedoch ein Verb ergänzt wurde. Beispiel:

4.

Tok: *Was für meine Silvesternacht*

TH1: *Was meine Silvesternacht angeht.*

Der obige PP-Kontext unter 4. ist insofern nicht interpretierbar, als die *für*-PP auf tok und das Äquivalent dieser Phrase auf TH1-Ebene nur in Bezug auf ein Verb interpretiert werden können. Das Verb fehlt jedoch auf tok-Ebene.

Achtung Kategorie „Rest“ (R):

PP-Kontexte, bei denen die auf TH1 gelöschte Präposition Bestandteil einer anderen PP auf tok ist, werden als „Rest“ annotiert. Beispiel:

5.

Tok: *Bitte helfen Sie mir [für bei der Wohnungssuchen].*

TH1: *Bitte helfen Sie mir [bei der Wohnungssuche .]*

-> Hier liegt auf tok-Ebene eine Dopplung von Präpositionen vor (*für* und *bei*), keine einfache Hinzufügung einer Präposition bzw. PP.

Weiterhin werden Kontexte, bei denen die PP eine Subjekt-Funktion erfüllt, ebenfalls als „Rest“ annotiert. Beispiel:

6.

Tok: *Aber [für die andere Seite] ist nicht immer schön*

TH1: *Aber [die andere Seite] ist nicht immer schön*

-> Hier wurde die PP auf tok-Ebene (*für die andere Seite*) auf TH1-Ebene in eine NP (*die andere Seite*) geändert. Diese NP erfüllt die Funktion des Subjekts im Satz. Subjekt-PP stehen in der Dissertation nicht im Fokus und werden daher in die Kategorie „Rest“ annotiert.

Weitere Beispiele für die Annotation der syntaktischen Funktion von PP des Fehlertyps „Hinzufügung“ finden sich im Dokument *Hinweise zur Annotation* (s. Anhang).

2.3 Weitere Annotationskategorien

2.3.1 Rest (R)

Folgende PP sollen in der Kategorie Rest annotiert werden:

1) PP, die aufgrund von Unsicherheiten in keine andere Kategorie klassifiziert werden können oder deren Zuordnung nicht eindeutig in eine bestimmte Kategorie vorgenommen werden kann; Grund dafür ggf. in der Kommentarspalte vermerken. Diese Fälle werden dann mit der Annotationsleitung diskutiert.

2) Instanzen der Fehlerkategorie „Hinzufügung“, bei denen die auf TH1 gelöschte Präposition Bestandteil einer anderen PP auf tok ist (s. oben): Beispiel:

Tok: *Bitte helfen Sie mir [für bei der Wohnungssuchen].*

TH1: *Bitte helfen Sie mir [bei der Wohnungssuche .]*

3) Instanzen der Fehlerkategorie „Hinzufügung“, bei denen die ‚ursprüngliche‘ PP (auf tok-Ebene) eine Subjekt-Funktion (TH1) erfüllt (s. oben). Beispiel:

Tok: *Aber [für die andere Seite] ist nicht immer schön*

TH1: *Aber [die andere Seite] ist nicht immer schön*

Hinweise zu der Annotation von PP, bei denen Präpositionen, die in der Original-Lerneräußerung (Ebene *tok*) enthalten sind, auf Ebene der TH1 jedoch gelöscht wurden, finden sich weiter oben (Kap. 2.2.5). Diese bitte aufmerksam lesen.

2.3.2 Pseudotreffer

Aufgrund der wortformbasierten Abfrage im Korpus kann es sein, dass unter den PP-Kontexten ebenfalls Pseudotreffer vorliegen. Beispiele (s. z. B. auch Duden 2016: 612):

- a) abtrennbares Verbpräfix: *Ich rufe dich an./Ich mache die Tür auf.*
- b) Postpositionen: *Der Meinung des Autors nach, ist die Gesellschaft [...]*
- c) Zirkumpositionen: *Von Beginn an; von Geburt an [...]*
- d) Präpositionsartige Fügungen bzw. tertiäre Präpositionen: *im Vergleich mit; in Bezug auf; an Stelle von; Im Verlauf(e) von*
- e) Komplexe Präpositionen (die z. T. noch getrennt geschrieben werden können): *mithilfe des Messers bzw. mit Hilfe des Messers, anstelle des Gerichts bzw. an Stelle des Gerichts*

Solche PP-Kontexte sind als „Pseudotreffer“ zu annotieren und werden in der weiteren Auswertung nicht berücksichtigt.

2.3.3 Nicht interpretierbar (NI)

Ein PP-Kontext ist als „nicht interpretierbar“ (NI) zu annotieren, wenn folgende Fälle vorliegen:
1) Wenn der Lenersatz (auf Ebene der TH1 (!)) aufgrund von lexikalischen, d. h. semantischen Abweichungen nicht interpretierbar bzw. unverständlich ist, da er z. B. Wörter enthält, die in der deutschen Sprache nicht existieren.

Beispiele:

- a) TH1: *Ich führe sich auf Sie* -> die Funktion der PP ist hier nicht interpretierbar, da das Verb „führe auf“ nicht interpretierbar ist.
- b) TH1: *1980 gab es in dieser 20 000 000 Eingewohner-Stadt eine Spitzung im Konstruktionsmarkt mit viele neue Gebäuden*. -> Die Interpretation der PP ist erschwert, weil das Wort *Spitzung* in der deutschen Sprache nicht existiert.

2) Tok-Ebene und TH1-Ebene stimmen nicht überein, d. h. der Satz auf tok-Ebene entspricht nicht dem Satz auf TH1-Ebene. Es handelt sich wahrscheinlich um einen Alinierungsfehler im Korpus.

3) Instanzen der Fehlerkategorie „Hinzufügung“, bei denen eine PP, die in der Original-Lerneräußerung (*tok*) verwendet wurde, auf Ebene der TH1 komplett gelöscht wurde bzw. es auf Ebene der TH1 keine Entsprechung für die PP gibt (s. auch weiter oben).

Beispiele:

1.

Tok: Du sollst mit dir ein Anzuge Und die Brille für Schwimmen mitbringen .

TH1: Du sollst einen Anzug und die Brille zum Schwimmen mitbringen .

-> Die PP *mit dir* auf der Tok-Ebene wurde auf Ebene der TH1 komplett gelöscht. D. h. es gibt auf TH1-Ebene keine PP-Entsprechung man für Annotationszwecke interpretieren kann.

2.

Tok: Ein Balkon auf .

TH1: , einen Balkon ____.

-> Hier liegt eine hinzugefügte Präposition auf tok-Ebene vor, auf der Ebene der TH1 findet man jedoch an dieser Stelle keine PP-Entsprechung. D. h. es gibt auf TH1-Ebene keine Phrase, die man für Annotationszwecke interpretieren kann.

Hinweise zu der Annotationen von PP, bei denen Präpositionen, die in der Original-Lerneräußerung (Ebene *tok*) enthalten sind, auf Ebene der TH1 jedoch gelöscht wurden, finden sich weiter oben (Kap. 2.2.5). Diese bitte aufmerksam lesen.

Achtung: Die Annotationskategorie „nicht interpretierbar“ nicht mit der Kategorie „Rest“ (s. u.) verwechseln. In die Kategorie „nicht interpretierbar“ werden Fälle einsortiert, die semantisch unverständlich bzw. für Annotationszwecke nicht interpretierbar sind. Die Kategorie „Rest“ enthält dahingegen vor allem Fälle, in denen man den Lerneratz zwar interpretieren kann, sich jedoch z. B. bezüglich der Zuordnung zu einer Kategorie unsicher ist.

2.3.4 Funktionsverbgefüge (FVG)

Als Funktionsverbgefüge (FVG) ist hier eine Konstruktion aus Funktionsverb (FV) und einer PP zu verstanden. Das FVG nimmt die Funktion eines komplexen Prädikats ein, da es in der Regel einem Vollverb entspricht (Duden 2016: 426, Eisenberg 2013: 205).

Nach Duden (2016: 429) zählen zu den FV, die mit einer PP kombiniert werden können, vor allem folgende Verben:

bringen, stellen, setzen, nehmen, ziehen, stehen, bleiben, halten.

Typische Präpositionen, die mit den FV in Verbindung treten sind: *zu/in*, andere Präpositionen wie *an/auf/unter/außer* sind jedoch auch möglich.

Beispiele für FVG:

zur Geltung kommen

zur Diskussion stehen

zur Anwendung bringen

in Vergessenheit geraten

zum Schmelzen bringen, etc.

(vgl. Duden 2016: 429f. , Eisenberg 2013: 305).

Weitere Eigenschaften von FVG mit Präpositionen sind (vgl. Eisenberg 2013: 308ff.):

(am Beispiel (*sich*) *in Form bringen*)

- Attribute zum Nomen nur begrenzt möglich (**sich in gute Form bringen*)
- Nomen ist nicht pronominalisierbar und nicht erfragbar (**Sich worin bringen?*)
- Nomen in der Regel im Singular (**sich in Formen bringen*)

- Negation mit „keine“ ausgeschlossen, nur Negation mit „nicht“ möglich (**sich in keine Form bringen*)
- Artikel liegen fest, sind nicht semantisch motiviert, oft als Verschmelzung (z. B. *zur Anwendung kommen*) oder ganz ausgeschlossen (z. B. *in Form bringen*).

2.3.5 Formelhafte Sequenz (FS)

Formelhafte Sequenzen (FS) sind Konstruktionen, die eine semantische Einheit darstellen, aber aus mehreren sprachlichen Einheiten (inkl. einer PP) aufgebaut sind. An dieser Stelle wird die Definition von *formulaic sequence*, die sich ebenfalls im MERLIN-Annotationsschema findet, angesetzt: “[...] a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (Wray 2002: 9).

Beispiel für PP-Kontexte, die „Formelhaften Sequenzen“ darstellen und als solche zu annotieren sind:

- *auf jeden/keinen Fall*
- Grußformeln: *Auf Wiedersehen! Mit freundlichen Grüßen*
- Idiome/idiomatische Konstruktionen (die etwas anderes bedeuten als ihre Bestandteile): *auf die Palme bringen; mit Sack und Pack; mit Kind und Kegel*
- *Wie wäre es mit X³¹⁶?/Wie ist es mit X?*

Weitere Beispiele für formelhafte Sequenzen sind im Dokument *Beispiele Annotationen* dokumentiert (s. Anhang).

Anmerkung: Da die Beschreibung dieser Kategorie nicht eindeutiger ausfallen kann, bitte nicht davor zögern, dahingehend unsichere Fälle wenn nötig in die Kategorie „Rest“ einzusortieren. Diese Fälle werden dann mit der Annotationsleitung diskutiert.

2.3.6 Kommentar

Neben den Annotationskategorien findet sich in jedem Annotationsdokument eine Kommentarspalte, in der Kommentare zu Annotation hinterlassen werden können bzw. hinterlassen werden müssen (z. B. Recherche in E-Valbu, s.o.). Unsicherheiten bei der Annotationen, Fragen etc. sollten in dieser Spalte notiert werden. Diese Fälle werden dann mit der Annotationsleitung diskutiert.

2.4 Annotationskategorien und –werte im Überblick

Es folgt ein tabellarischer Überblick über die einzelnen Annotationskategorien und entsprechende Annotationsbeispiele (Tab. 1).

³¹⁶ Das „X“ gilt hier nur als Platzhalter. Verschiedene Ausdrücke sind an dieser Stelle möglich. Unabhängig vom Ausdruck werden solche Konstruktionen als „formelhafte Sequenz“ annotiert.

Hinweis: Pro Kategorie bitte jeweils 0 bzw. 1 in die entsprechende Spalte eintragen, leere Spalten vermeiden.

Annotationskategorie	Erklärung/Beispiele	Annotationswerte ³¹⁷
Syntaktische Funktion der PP Objekt-PP (O) (Hinweis: E-Valbu o. Duden konsultieren und Lesart im Kommentar vermerken)	PP=Objekt , P ³¹⁸ von Verb regiert. Beispiele: <i>interessieren für</i> <i>freuen auf/über</i> <i>warten auf</i> <i>sprechen mit</i>	1/0
Syntaktische Funktion der PP Objekt-PP Adjektiv (OAdj) (Hinweis: Duden konsultieren und Lesart im Kommentar vermerken)	PP= von einem Adjektiv regiert. Beispiele: <i>stolz auf den Erfolg</i> <i>interessiert an der Arbeit</i>	1/0
Syntaktische Funktion der PP Adverbiale PP (A)	PP=Adverbiale. P i.d.R. nicht vom Verb regiert. Beispiele: <i>Ich esse auf/neben/bei der Parkbank</i>	1/0
Syntaktische Funktion der PP Attributive PP (Att)	PP= Attribut. PP ist kein eigenständiges Satzglied, sondern Gliedteil; i.d.R. ist PP nicht verschiebbar. Beispiele: <i>Die Freude auf den Urlaub</i> <i>Ein Lehrer mit reichlich Erfahrung</i> <i>Eine große Anzahl an [...] Rechnern</i>	1/0

³¹⁷ 1= ja/liegt vor; 0 = nein/liegt nicht vor. Im Annotationsdokument die Annotationswerte bitte auf Höhe der TH1, d. h. in Höhe der TH1-Zeile, in die jeweiligen Spalten eintragen.

³¹⁸ P steht für Präposition.

Rest (R)	<ul style="list-style-type: none"> - unsichere Fälle bzw. PP, die nicht eindeutig einer anderen Kategorie zugeordnet werden können. - Instanzen von PP des Fehlertyps <i>Hinzufügung</i>, bei denen die PP die Funktion eines Subjekts einnimmt. Beispiel: Tok: <i>Aber <u>für die andere Seite</u> ist nicht immer schön</i> TH1: <i>Aber <u>die andere Seite</u> ist nicht immer schön</i> 	1/0
Pseudotreffer (P)	<p>Beispiele:</p> <ul style="list-style-type: none"> - abtrennbares Verbpräfix (<i>Ich rufe dich <u>an</u>.</i>) - präpositionsartige Fügung: <i>in Bezug auf/im Vergleich mit etc.</i> - komplexe Präposition: <i>mithilfe/mit Hilfe etc.</i> 	1/0
Nicht interpretierbar (NI)	<ul style="list-style-type: none"> - Kontext auf TH1-Ebene (!) unverständlich und kann nicht weiter in Bezug auf die PP analysiert/interpretiert werden. - Instanzen von PP des Fehlertyps <i>Hinzufügung</i>, bei denen es auf der TH1-Ebene keine Entsprechung zur PP auf der tok-Ebene gibt. <p>Beispiel: Tok: <i>Ein Balkon <u>auf</u> .</i> TH1: <i>, einen Balkon ____.</i></p>	1/0
Funktionsverbgefüge (FVG)	<p>Die PP ist Bestandteil eines FVG aus einem Funktionsverb und PP.</p> <p>Beispiele: <i>In Verbindung bringen, In Vergessenheit geraten, Zur Anwendung kommen</i></p>	1/0
Formelhafte Sequenz (FS)	<p>Die PP ist Teil einer formelhaften Sequenz, die aus mehreren sprachlichen Einheiten besteht. Dabei kann es sich auch um ein Idiom handeln. Beispiele: <i>auf jeden Fall</i> <i>Auf Wiedersehen, Auf die Palme bringen</i> <i>Mit Sack und Pack, Wie ist/wäre es mit X?</i></p>	1/0

Tabelle 1: Annotationskategorien und -werte: Zusammenfassung mit Beispielen.

2.4.1. Ausgewählte Annotationsbeispiele

Die folgende Tabelle (Tab. 2) enthält ausgewählte Annotationsbeispiele zur Annotation der syntaktischen Funktion der PP. Weitere Annotationsbeispiele (auch für die anderen Kategorien) finden sich im Dokument *Beispiele Annotationen* (s. Anhang).

Hinweis für Annotation im Excel-Dokument:

Im Annotationsdokument die Annotationswerte bitte auf Höhe der TH1 in die jeweiligen Spalten eintragen, d. h. in Höhe der TH1-Zeile.

Beispiel	O	OAdj	A	Att
<p>Tok: ich warte <u>Ihre positive Antworten</u></p> <p>TH1: und ich warte <u>auf Ihre positiven Antworten</u></p> <p>Anmerkung: Verb regiert PP.</p>	1	0	0	0
<p>Tok: auf der Suche nach einem lehrreichen und anspruchsvollen Arbeitsumfeld bin ich <u>auf ihr Unternehmen</u> aufmerksam geworden</p> <p>TH1: auf der Suche nach einem lehrreichen und anspruchsvollen Arbeitsumfeld bin <u>auf Ihr Unternehmen</u> aufmerksam geworden</p> <p>Anmerkung: Adjektiv regiert PP.</p>	0	1	0	0
<p>Tok: [...] ich bin <u>für diese Stelle</u> sehr interessiert.</p> <p>TH1: [...] ich bin <u>an dieser Stelle</u> interessiert.</p> <p>Anmerkung: Adjektiv regiert PP.</p>	0	1	0	0
<p>Tok: Seit meiner Kindheit hatte ich das Traum, <u>in einem Au-pair Programm</u> in Deutschland teilzunehmen.</p> <p>TH1: Seit meiner Kindheit hatte ich den Traum, <u>an einem Au-pair Programm</u> in Deutschland teilzunehmen.</p> <p>Anmerkung: Verb regiert PP.</p>	1	0	0	0
<p>Tok: <u>In diesen Plätzen</u> können wir die ungarische Kultur erfahren.</p> <p>TH1: <u>An diesen Plätzen</u> können wir die ungarische Kultur erfahren.</p> <p>Anmerkung: Verb regiert PP nicht; PP ist weglassbar; PP stellt eine modale Spezifizierung dar.</p>	0	0	1	0
<p>Tok: weil die andere Kulturen gute Auswirkung <u>auf die Gesellschaft</u> des Gastlandes haben</p> <p>TH1: weil die anderen Kulturen eine gute Auswirkung <u>auf die Gesellschaft</u> des Gastlandes haben</p> <p>Anmerkung: PP ist Attribut zu „Auswirkung“.</p>	0	0	0	1

Tok: Ich habe <u>am bald</u> Gebutstag [...]				
TH1: Ich habe <u>bald</u> Geburtstag [...]				
Anmerkung: Hier ist <i>am bald</i> (tok) zu <i>bald</i> (TH1) geändert worden. Auf TH1 hat <i>bald</i> die Funktion eines Adverbiale.	0	0	1	0

Tabelle 2: Ausgewählte Annotationsbeispiele.

Beispiel Auszug aus Annotationsdokument (GeR-Niveau B2, PP mit *an*)

Niveau B2 Präposition: an			Fehlertyp	O Objekt-PP	A Adverbiale PP	Att Attributive PP	OAdj Objekt-PP Adjektiv	R Rest	Pseudot
0.	tok	ich habe Ihre Anansee in der Zeitung gelesen . Von ihr Anzeige bin ich interessiert . Für diese Stelle bringe							
	TH1	, ich habe Ihre Annonce in der Zeitung gelesen . An ihrer Anzeige bin ich interessiert . Für diese Stelle bringe	ch	0	0	0	1		0
	G_Prep_type	ch							
	meta::_author_id	1023_0001418							
	meta::_author_L1	other							

Anmerkungen zu Abkürzungen im Excel-Annotationsdokument:

tok= Original-Lerneräußerung

TH1= Zielhypothese 1 (enthält ggf. Korrekturen im Bereich Grammatik und Orthographie)

G_Prep_type=Fehlertyp im Bereich Präpositionsgebrauch (nicht immer vorliegend) (Hinweise zur Annotation s. unten).

meta::_author_L1= Erstsprache der Lernenden

meta::_author_id= Identifikationsnummer des Lernertexts

Die Annotationskategorien sind in Spalten nebeneinander angeordnet.³¹⁹ Im oberen Beispiel liegt eine PP (TH1: *an Ihrer Anzeige*³²⁰) vor, die von einem Adjektiv (hier: *interessiert*) regiert wird. Die PP wird folglich in die Kategorie „OAdj“ (Objekt-PP Adjektiv) annotiert (Annotationswert 1 in dieser Kategorie). Die anderen Spalten bzw. Kategorien erhalten den Wert 0. Die Annotationswerte werden auf Ebene der TH1 (= gleiche Zeile wie TH1) eingetragen (s. Auszug aus Annotationsdokument oben).

2.5 Annotation Fehlertyp im Bereich Präpositionsgebrauch

Im Falle eines zielsprachlich abweichenden Gebrauchs von Präpositionen bzw. PP wird der Fehler (bzw. der Fehlertyp) im Bereich Präpositionsgebrauch annotiert. Z.T. liegen in den exportierten PP-Kontexten bereits Fehlerannotationen im Bereich Präpositionsgebrauch vor, auf die bei der Annotation des Fehlertyps im Annotationsdokument Bezug genommen werden soll.

³¹⁹ In den Annotationsdokumenten sind die Spalten bunt eingefärbt. Für die Darstellung im Annotationshandbuch wird jedoch auf diese bunte Einfärbung verzichtet.

³²⁰ Die Original-Lerneräußerung (tok) enthält hier die PP *von ihrer Anzeige* (s. tok). Die Annotation der Fehler im Bereich Präpositionsgebrauch wird im Handbuch weiter unten beschrieben.

In diesem Kapitel wird das Vorgehen bei der Annotation des Fehlertyps im Bereich Präpositionsgebrauch beschrieben. Zuerst werden allgemeine Hinweise zur Annotation des Fehlertyps gegeben und die Annotationskategorien und -werte beschrieben. Im Anschluss folgen Hinweise zum Umgang mit fehlenden Fehlerannotationen sowie zum Umgang mit inkorrekten Fehlerannotationen.

Die Hinweise in diesem Kapitel sind nur für die PP-Kontexte relevant, die als *Objekt-PP* (bzw. *Objekt-PP Adjektiv*), *adverbiale PP* oder *attributive PP* annotiert wurden. Bei PP-Kontexten, die als „Rest“ annotiert wurden, ist die Kategorie „Fehlertyp“ zunächst frei zu lassen. Sollten PP aus der Kategorie „Rest“ nach Abschluss der Diskussionen mit der Annotationsleitung eine andere Annotation erhalten, dann wird die Annotation des Fehlertyps für diese PP-Kontexte nachgetragen.

2.5.1 Annotation Fehlertyp: Überblick

Tabelle 3 gibt einen ersten Überblick über die Annotationskategorien und –werte zur Annotation des Fehlertyps im Präpositionsgebrauch.

Annotationskategorie Fehlertyp im Bereich Präpositionsgebrauch	
Erläuterung (Fehlertyp)	Annotationswert 321
Kein Fehler Es liegt kein Fehler im Bereich Präpositionsgebrauch vor.	0 (=null)
Auslassung: Präposition fehlt auf tok und wurde auf TH1 ergänzt.	o (omission)
Wahl: Präposition auf tok wurde auf TH1 geändert.	ch (choice)
Position: Position der Präposition auf tok wurde auf TH1 geändert.	pos (position)
Hinzufügung: Präposition wurde auf tok hinzugefügt und auf TH1 gelöscht.	ad (addition)
Zusammenziehung: Präposition wurde auf tok falsch verschmolzen (mit einem anderen Wort), Verschmelzung wurde auf TH1 aufgehoben.	merge
Trennung: Präposition wurde auf tok falsch getrennt geschrieben, Trennung wurde auf TH1 aufgehoben.	split

Tabelle 3: Annotationskategorien und -werte zur Annotation des Fehlertyps im Bereich Präpositionsgebrauch: Überblick.

³²¹ Im Annotationsdokument bitte nur den Wert eintragen, der in dieser Tabellenspalte nicht im Klammern steht, also bei dem Fehlertyp *Wahl* nur den Wert *ch* eintragen.

Weitere Hinweise und Beispiele zu den Kategorien finden sich im MERLIN-Annotationsschema (MERLIN project 2014: 12)³²².

2.5.2 Allgemeine Hinweise zur Annotation des Fehlertyps im Präpositionsgebrauch

Bei den Korpustreffern bzw. PP-Kontexten in tok, in denen ein Fehler im Bereich Präpositionsgebrauch vorliegt, wurden die Fehlerannotationen ebenfalls aus dem Lernerkorpus extrahiert (die entsprechende Zeile im Annotationsdokument lautet: *G_Prep_type*). In der Regel gibt es bei PP, die keine Fehler enthalten, auch keine Annotation des Fehlertyps, d. h. die Zeile *G_Prep_type* fehlt entsprechend.³²³

Vorhandene Annotationen zum Fehlertyp im Präpositionsgebrauch (Kürzel *o*, *ch*, *pos*, *ad*, *merge*, *split* in der Zeile *G_Prep_type*) sollen im Annotationsdokument in die Spalte *Fehlertyp* eingetragen werden, s. Beispiel Auszug aus dem Annotationsdokument (Abb. 9):

Niveau B2 Präposition: an			Fehlertyp	O Objekt-PP	A Adverbiale PP	Att Attributive PP	OAdj Objekt-PP Adjektiv	R Rest
0.	tok	ich habe Ihre Anannse in der Zeitung gelesen . Von ihr Anzeige bin ich interessiert . Für diese Stelle bringe						
	TH1	, ich habe Ihre Annonce in der Zeitung gelesen . An ihrer Anzeige bin ich interessiert . Für diese Stelle bringe	ch	0	0	0	1	
	<i>G_Prep_type</i>	ch						
	meta::_author_id	1023_0001418						
	meta::_author_L1	other						

Abbildung 9: Ausschnitt aus Annotationsdokument: Annotation Fehlertyp (1): Beispiel.

Zum Ausschnitt aus dem Annotationsdokument (Abb. 9) :

In der Zeile *G_Prep_type* (unter TH1-Zeile) findet sich ein Kürzel für den Fehlertyp im Bereich Präpositionsgebrauch, hier: *ch*. Dieses Kürzel wird in die Spalte *Fehlertyp* eingetragen (s. Abb. 9).

Anmerkung: Manchmal kommt es vor, dass in dem exportierten PP-Kontext mehrere Fehler im Bereich Präpositionsgebrauch vorliegen. Dementsprechend finden sich auch mehrere Werte (= Kürzel) in der Zeile *G_Prep_type*. Dennoch entspricht jeder exportierte (und im Annotationsdokument nummerierte) PP-Kontext einem (!) PP-Kontext, folglich ist je exportierten PP-Kontext auch max. ein Fehler im Bereich Präpositionsgebrauch zu annotieren. D. h. in einer *G_Prep_type*-Zeile, die mehrere Werte enthält, ist der Wert zu bestimmen, der zum aktuell annotierten PP-Kontext gehört. Dieser Wert wird dann in die Annotationsspalte *Fehlertyp* eingetragen. Siehe dazu das folgende Beispiel in Abbildung 10.

³²² Das Dokument ist online frei verfügbar auf <http://www.merlin-platform.eu>

³²³ Für abweichende Fälle, in denen trotz eines Fehlers im Präpositionsgebrauch keine Fehlerannotation vorliegt, siehe weiter unten.

Niveau B2 Präposition: an			Fehlertyp	O	Objekt-PP	A	Adverbiale PP	Att	Attributive PP	Objl	Objekt-PP Adjektiv	R	Res
141.	tok	Aber man muß auch an der Kultur des Gastlandes kennenzulernen . Wenn man nach Ausland											
	TH1	Aber man muss auch die Kultur des Gastlandes kennenlernen . Wenn man ins Ausland	ad		1		0		0		0		
	G_Prep_type	ad ch											
	meta::rating_fair	B2											
	meta::author_id	1031_0003231											

Abbildung 10: Beispiel Ausschnitt aus Annotationsdokument: Annotation Fehlertyp (2): Beispiel.

Zum Ausschnitt Annotationsdokument (Abb. 10) :

In tok finden sich zwei Fehler im Bereich Präpositionsgebrauch, es sind folglich zwei Werte in der Zeile *G_Prep_type* zu finden: *ad* und *ch*. Der hier zu annotierende Kontext betrifft die PP mit *an* (hier: *an der Kultur*). Diese Präposition ist auf tok hinzugefügt, wurde auf Ebene der TH1 jedoch gelöscht (*die Kultur*) (= Fehlertyp *ad*). Folglich wird in der Spalte *Fehlertyp* der Wert *ad* annotiert (s. Tab. 3).

Es kann ebenfalls vorkommen, dass im Datenausschnitt Fehler im Präpositionsgebrauch vorliegen, diese jedoch nicht die Präposition bzw. PP betreffen, die aktuell annotiert werden soll, vgl. das folgende Beispiel in Abbildung 11.

Präpositin mit			Fehlertyp	O	Objekt-PP	A	Adverbiale PP	Att	Attributive PP	Komn
27.	tok	Schulabschluss mit sehr guten Noten habe ich begonnen Informatic Schwerpunkt mit Systemkaufmann , ...									
	TH1	dem Schulabschluss mit sehr guten Noten habe ich Informatik mit Schwerpunkt Systemkaufmann begonnen , ...		0	0		0		1	...	
	G_Prep_type	pos									
	meta::author_id	1023_0101894									
	meta::author_L1	not reported									

Abbildung 11: Ausschnitt aus Annotationsdokument: Annotation Fehlertyp (3): Beispiel.

Zum Ausschnitt Annotationsdokument (Abb. 11):

Auf tok findet man zwei *mit*-PP, wovon hier der PP-Kontext der ersten *mit*-PP (*mit sehr guten Noten*) annotiert werden soll. Im Falle dieser *mit*-PP liegt kein Fehler im Bereich Präpositionsgebrauch vor (vgl. TH1: *mit sehr guten Noten*); im Falle der zweiten *mit*-PP (tok: *Informatic Schwerpunkt mit Systemkaufmann*, TH1: [...] *mit Schwerpunkt Systemkaufmann*) wurde die Position der Präposition auf tok-Ebene auf der TH1-Ebene geändert, hier liegt folglich der Fehlertyp Position (Kürzel *pos*) vor. In der Zeile *G_Prep_type* findet man das Kürzel dieses Fehlertyps. Da in diesem Beispiel jedoch die Annotation der erstgenannten *mit*-PP im Fokus steht, die keinen Fehler im Bereich Präpositionsgebrauch enthält, wird für diese *mit*-PP in der Spalte *Fehlertyp* entsprechend 0 (=kein Fehler) eingetragen (s. Abb. 11).

Die obigen Beispiele zeigen: Es ist notwendig, die tok-Ebene und die vorliegenden Werte in der Zeile *G_Prep_type* genau zu studieren, um zu bestimmen, ob die Werte den jeweils zu annotierenden PP-Kontext betreffen bzw. welche Werte den Fehlertypen des zu annotierenden PP-Kontexts abbilden.

2.5.3 Annotation Fehlertyp: Umgang mit fehlenden Fehlerannotationen

Wird bei der Annotation festgestellt, dass eine Fehlerannotation im Bereich Präpositionsgebrauch fehlt, d. h. die Zeile *G_Prep_type* fehlt, obwohl ein Fehler im Präpositionsgebrauch des zu annotierenden PP-Kontexts vorliegt, wird der Fehlertyp manuell nachannotiert. Das entsprechende Vorgehen wird hier beschrieben.

Wichtig: Als Bezugsebene TH1 beachten!

Bei der Annotation fehlender Fehlerannotationen muss man sich an der TH1-Ebene und den Änderungen dort orientieren. Beispiel: Wurde auf der TH1 eine Präposition gelöscht, die auf tok vorhanden ist, so ist in der Spalte *Fehlertyp* entsprechend ein *ad* einzutragen (vgl. Tab. 3). Bei der Annotation in der Spalte *Fehlertyp* ist wie unter 2.5.2 beschrieben vorzugehen und den Annotationswert für den entsprechenden Fehlertyp zu wählen.

In der Kommentarspalte dokumentiert man die manuelle Korrektur durch eine Kombination aus dem Kürzel TW und dem Annotationswert des Fehlertyps (ohne Leerzeichen). Beispiel: „TWch“ (für eine Annotation des Fehlertyps *ch*, die manuell ergänzt wurde), s. die folgende Abbildung (Abb. 12):

Niveau B2 Präposition: an			Fehlertyp	O	Objekt-PP	A	Adverbiale PP	Att	Attributive PP	Kommentar
54.	tok	. Ich hatte überhaupt nicht keine Freizeit Ich wollte in eine Deutschschule Teilnehme aber ich habe vollzeit gearbeitet .									
	TH1	. Ich hatte überhaupt keine Freizeit . Ich wollte an einer Deutschschule teilnehmen , aber ich habe Vollzeit gearbeitet .	ch		1		0		0		TWch
	meta::_author_id	1023_0109249									
	meta::_author_L1	Spanish									

Abbildung 12: Annotationsdokument: Annotation Fehlertyp (fehlende Fehlerannotation): Beispiel.

Zum Ausschnitt Annotationsdokument (Abb. 12):

Die PP *in eine Deutschschule Teilnehme* (tok) wurde auf TH1 geändert zu *an einer Deutschschule teilnehmen*. Hier liegt demnach ein Fehler im Bereich Präpositionsgebrauch vor, und zwar der Fehlertyp *Wahl* (Annotationswert *ch*). Der entsprechende Annotationswert wird nun in die Spalte *Fehlertyp* manuell eingetragen und die Annotation wird in der Kommentarspalte mit *TWch* dokumentiert.

2.5.4 Annotation Fehlertyp: Umgang mit inkorrekten Fehlerannotationen

Wird bei der Annotation festgestellt, dass eine Fehlerannotation im Bereich Präpositionsgebrauch, die sich eindeutig auf den zu annotierenden PP-Kontext bezieht, inkorrekt ist, d. h. die Zeile *G_Prep_type* enthält zwar einen Wert, dieser ist aber inkorrekt, so wird dieser Fehler korrigiert und in der Spalte *Fehlertyp* manuell nachannotiert. Dabei ist bei der Annotation in der Spalte *Fehlertyp* wie unter 2.1. beschrieben vorzugehen und das entsprechende Kürzel (vgl. Tab. 3) für den korrekten Fehlertyp zu wählen.

Wichtig: Als Bezugsebene TH1 beachten! Bei der Korrektur der inkorrekten Fehlerannotationen muss man sich an der TH1-Ebene und den Änderungen dort orientieren.

Die Korrektur inkorrekt Fehlerannotationen wird in der Kommentarspalte dokumentiert, und zwar durch eine Kombination aus dem Kürzel *corrTW* und dem

Annotationswert des Fehlertyps (ohne Leerzeichen), also *corrTWxx* , wobei *xx* auf den korrekten Fehlertyp verweist. Beispiel: *corrTWch*, bedeutet: Der extrahierte Fehlertyp war inkorrekt und wurde im Rahmen der Annotation zum Fehlertyp *ch* korrigiert, vgl. das Beispiel in Abbildung 13:

Niveau B1 Präposition: auf			Fehlertyp	O	Objekt-PP	A	Adverbiale PP	Att.	Attributive PP	Kommentar
51.	tok	Ich kann leider nicht fahren , ob du lust hast nach Frankfurt , Du kannst das Ticket nehmen , es ist									
	TH1	Ich kann leider nicht fahren ; ob du Lust hast auf Frankfurt ? Du kannst das Ticket nehmen , es ist	ch		0		0			1	corrTWch
	G_Prep_type	o									
	meta::_author_l1	not reported									
	meta::_author_id	1091_0000002									

Abbildung 13: Annotationsdokument: Annotation Fehlertyp (inkorrekte Fehlerannotation): Beispiel.

Zum Ausschnitt Annotationsdokument (Abb. 13):

Auf tok-Ebene findet sich die PP *lust [...] nach Frankfurt*, auf TH1-Ebene wurde diese PP korrigiert zu *Lust [...] auf Frankfurt*, d. h. die Präposition in der PP wurde geändert. Die Fehlerannotation enthält jedoch den Annotationswert zum Fehlertyp *Auslassung (o)*. Dieser Annotationswert wird korrigiert und der korrigierte Wert in der Spalte *Fehlertyp* eingetragen (Annotationswert *ch* für Fehlertyp *Wahl*), die Korrektur wird in der Kommentarspalte mit *corrTWch* dokumentiert.

Für weitere Hinweise zu Fehlerannotationen im Bereich Präpositionsgebrauch kann das Annotationsschema des Lernerkorpus MERLIN konsultiert werden (Wisniewski et al. 2014: 12). Das Annotationsschema ist auf der MERLIN Plattform³²⁴ frei verfügbar.

Für die Auswertung der Annotationen werden lediglich die Fehlertypen in der Annotationsspalte *Fehlertyp* berücksichtigt.

2.6 Annotation Gewähltes Verb/Adjektiv

Bei PP-Kontexten, die als *Objekt-PP* bzw. *Objekt-PP Adjektiv* annotiert werden, sollen zusätzlich die gewählten Verben bzw. Adjektive, die die jeweiligen PP regieren, in einer separaten Annotation erfasst werden. Dabei spielt es keine Rolle, ob ein Fehler im Bereich Präpositionsgebrauch vorliegt oder nicht. Das gewählte Verb bitte in Kleinbuchstaben im Infinitiv in die Annotationsspalte „Gewähltes V/Adj.“ eintragen, s. dazu folgende Ausschnitte aus Annotationsdokumenten (Abb. 14 und Abb. 15).

³²⁴ https://merlin-platform.eu/C_download.php.

Niveau B1 Präpositin für			FehlerTyp	O	Objekt-PP	A	Adverbale PP	Att	Attributive PP	...	Gewähltes V/Adj
0.	tok	möchte gern Teilnehmen . Weil habe ich intersieren mich für diesen fach und möchte gern Teil									
	TH1	möchte gern teilnehmen . Weil ich mich für dieses Fach interessiert habe , möchte ich gern teilnehmen	0	1	0	0	0	0			interessieren
	meta::_author_id	1023_0101689									
	meta::_author_L1	not reported									

Abbildung 14: Annotation gewähltes Verb/Adjektiv: Beispiel (1).

Niveau B2 Präpositin auf			FehlerTyp	O	Objekt-PP	A	Adverbale PP	Att	Attributive PP	...	Objekt-PP Adjektiv	Gewähltes V/Adj
148.	tok	mir Bescheid , wenn es geht . Ich bin gespannt auf deine Antwort . Herzliche Grüße Katharina										
	TH1	mir Bescheid , wenn es geht . Ich bin gespannt auf deine Antwort . Herzliche Grüße Katharina	0	0	0	0	0	0	0	0	1	gespannt
	meta::_author_id	1061_0120481										
	meta::_author_L1	English										

Abbildung 15: Annotation gewähltes Verb/Adjektiv: Beispiel (2).

3. Ergänzende Hinweise zur Nachrecherche im Lernerkorpus MERLIN

Sollten die exportierten PP-Kontexte im Annotationsdokument nicht umfangreich genug sein, um die PP eindeutig annotieren zu können, so empfiehlt es sich, mithilfe der entsprechenden Identifikationsnummer (Zeile *meta::_author_id*) direkt im Lernerkorpus MERLIN nach dem jeweiligen PP-Kontext zu suchen. Dafür geht man folgendermaßen vor:

1. Aufruf der Seite <https://commul.eurac.edu/annis/merlin> (direkter Korpus-Zugang)
2. Links mittig unter *Corpus List* Auswahl treffen: *Korpus MERLIN_German*.

In das obere Suchfenster folgende Anfrage eintragen: *meta::_author_id="ID" & tok* („ID“ steht hier für die entsprechende Identifikationsnummer des Lernertexts, z. B. 1023_0101688).
Beispiel: *meta::_author_id="1023_0101688" & tok*. Anschließend auf *Search* klicken (vgl. Abb. 16).

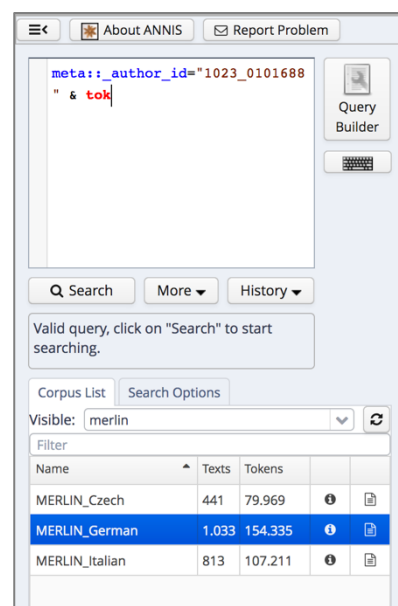


Abbildung 16: Abfrage im Lernerkorpus MERLIN: Beispiel (Screenshot).

3. Das Ergebnis der Abfrage wird auf der rechten Seite angezeigt. In dieser Übersicht auf der rechten Seite kann man dann Ausschnitte aus der Original-Lerneräußerung sehen, inkl. der verschiedenen Annotationsebenen (s. Abb. 17).

The screenshot shows a web interface for the MERLIN corpus. The path is 'MERLIN_German > 1023_0101688 (learner 1 - 6)'. There are dropdown menus for 'left context: 5' and 'right context: 5'. Below this is a grid for 'TH/EA (grid)' with columns for 'learner', 'Julia', 'Schmidt', 'Bahnhofstr', '.', '123', and '12345'. The 'TH1' row is highlighted. Below the grid are several annotation options: 'transcript (grid)', 'automatic (grid)', 'dependencies (arcs)', and 'full text'.

learner	Julia	Schmidt	Bahnhofstr	.	123	12345
TH1	Julia	Schmidt	Bahnhofstr	.	123	12345
TH1Diff			CHA			

Abbildung 17: Beispiel für einen Treffer im Lernerkorpus MERLIN. Links unten die Funktion +full text, rechts oben die Funktionen left context und right context (Screenshot).

Unter der Funktion *full text* kann der vollständige Lernertext eingesehen werden (dafür auf das +-Symbol drücken). Zur Erweiterung des Kontexts auf Ebene der TH1 können die Knöpfe „left context“ und „right context“ (jeweils oben rechts bei jedem Treffer, der in der Übersicht angezeigt wird) genutzt werden, um den Kontext bzw. die Anzahl der angezeigten Wörter rechts und links zu verändern. Dafür die entsprechenden Einstellungen im Drop-Down Menü bei „left context“ und „right context“ wählen (s. Abb. 17).

4. Im Anschluss den recherchierten erweiterten PP-Kontext bitte in das entsprechende Annotationsdokument eintragen, und zwar direkt in den Text in der Zeile *tok* und *TH1*. Das, was ergänzt wird, bitte in roter Farbe markieren.

Bei Unsicherheiten in Bezug auf die Recherche im Korpus den PP-Kontext in die Kategorie „Rest“ einsortieren und in der Kommentarspalte *Kontext* notieren. Diese Fälle werden dann mit der Annotationsleitung diskutiert.

4. Annotationen: Abschließende Hinweise und Checkliste

Bei den Annotationen ist es sehr wichtig, konsistent zu arbeiten. D. h. ähnliche PP sollten auch ähnlich annotiert sein, hierzu können beispielsweise persönliche Notizen, die man während der Annotation anfertigt, sehr hilfreich sein. Am Ende der Annotationen müssen die Annotationen daher immer auch durch die Annotatorin auf interne Konsistenz überprüft werden. Zusätzlich zur Konsistenz sollten die Annotationen auch immer auf Vollständigkeit hin überprüft werden; hierzu dient der Überblick über die Annotationsschritte (a) sowie die Checkliste zur Annotation (b), die im Folgenden dargestellt sind.

a) Überblick zu Annotationskategorien und einzelnen Annotationsschritten:

Die folgende Abbildung (Abb. 18) fasst die Annotationskategorien nochmal zusammen und macht die einzelnen Schritte im Annotationsprozess deutlich.

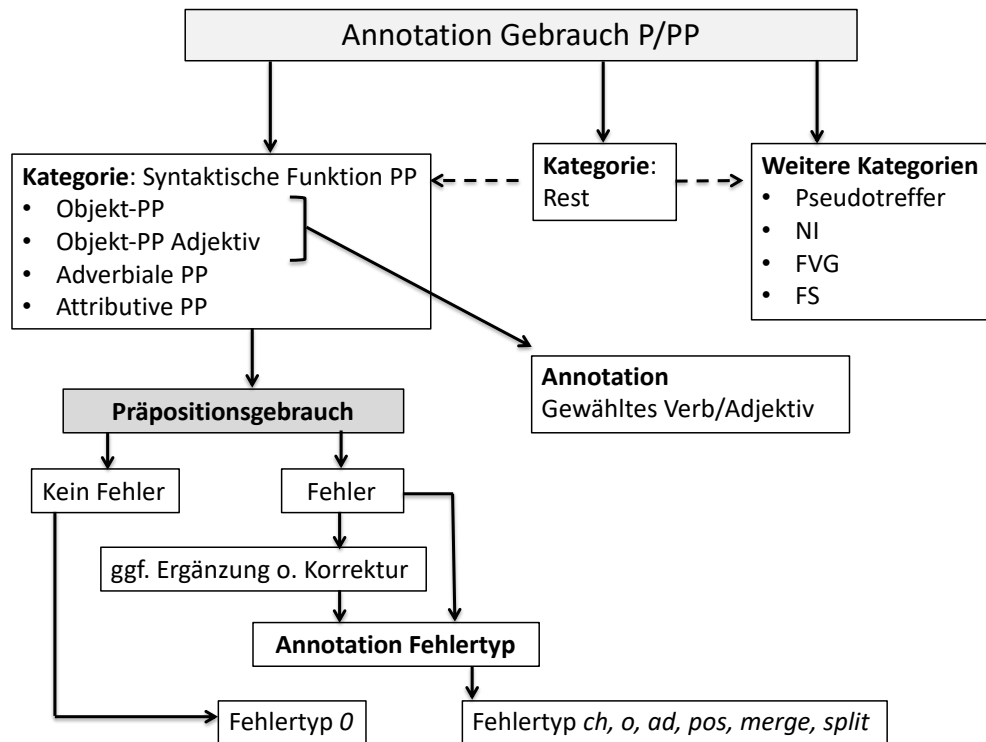


Abbildung 18: Annotationskategorien und Annotationsschritte: Zusammenfassung.

b) Checkliste Annotation:

Diese Checkliste soll während des Annotationsprozesses (und vor allem bei Abschluss der Annotationen) unbedingt herangezogen werden, vor allem um formale Annotationsfehler³²⁵ zu vermeiden. Die Checkliste erhebt keinen Anspruch auf Vollständigkeit und kann gerne durch die Annotatorin erweitert werden.

- Sind alle PP-Kontexte annotiert worden?
- Wurde E-Valbu/Duden konsultiert (und dies dokumentiert)?
- Wurde bei Objekt-PP die entsprechende Lesart des Verbs dokumentiert?
- Wurde das Dokument *Beispiele Annotationen* und die darin enthaltenen Annotationsbeispiele berücksichtigt?
- Wurde die Liste der Verben mit Objekt-PP/Objekt-PP Adjektiven beachtet? (s. Kap. 2.2.1)

³²⁵ Unter formalen Annotationsfehlern verstehe ich beispielsweise die irrtümliche Annotation eines PP-Kontexts in zwei unterschiedliche Annotationskategorien (z. B. bei einem PP-Kontext findet sich der Annotationswert 1 sowohl bei der Kategorie *Objekt-PP* als auch bei *attributiven PP*) oder die Annotation des gewählten Verbs/Adjektivs bei PP, die nicht als *Objekt-PP* bzw. *Objekt-PP Adjektiv* annotiert sind. Derartige Annotationen stellen Annotationsfehler dar, die es unbedingt zu vermeiden gilt.

- Wurden die Fehlertypen im Bereich Präpositionsgebrauch, so wie im jeweiligen PP-Kontext vorliegend, annotiert?
- Wurde bei der Dokumentation der Korrekturen bzw. Ergänzungen der Fehlertypen im Präpositionsgebrauch darauf geachtet, ob die Kürzel in der Zeile *G_Prep_type* die zu annotierende Präposition betreffen?
- Wurden notwendige Korrekturen bzw. Ergänzungen der Fehlertypen im Präpositionsgebrauch entsprechend dokumentiert?
- Wurden bei der Kategorie ‚Gewähltes Verb/Adjektiv‘ nur PP-Kontexten der Kategorien *Objekt-PP* und *Objekt-PP Adjektiv* berücksichtigt?
- Wurde bei PP, die als *attributive PP* annotiert wurden, der Verschiebe-Test angewandt?
- Sind die Annotationen in sich konsistent? D. h. wurden ähnliche PP auch entsprechend ähnlich annotiert?
- ...

5. Literaturverzeichnis und Ressourcen

s. Literaturverzeichnis der Dissertation

Anhang zum Annotationshandbuch

- Dokument *Beispiele Annotationen* (Fallstudie 3)
- Dokument *Beispiele zur Annotation von PP-Kontexten des Fehlertyps Hinzufügung*

Dokument *Beispiele Annotationen* (Fallstudie 3)

In dieser Tabelle finden sich Beispiele für die Annotation von PP mit <i>mit/für</i> . Stand: Juli 2018.		
Annotationskategorie	Beispiele für PP-Kontexte	
Objekt-PP	<p>Verben mit PP als Objekt (es sind natürlich immer die Lesart zu berücksichtigen, daher ist immer eine Kontrolle in E-Valbu durchzuführen; Hinweise auf Lesarten beziehen sich auf E-Valbu)</p> <p>anfangen [mit etw.] arbeiten [für etw.] (sich für etw. einsetzen) (E-Valbu, Lesart 5) sich bedanken [für XY] -> für-PP ist laut e-Valbu fakultativ, aber es handelt sich dennoch um eine Verb-Ergänzung bedeuten [für etw./jmdn.] (Lesart 3) beginnen [mit etw.] bekanntmachen [mit etw.] sich beschäftigen [mit etw.] (sich beschäftigen, Lesart 1) (sich) bewerben [für etw.] (Lesart 4) bezahlen [für etwas] (Lesart 2) brechen [mit jdm.] danken [für etw.] empfehlen [für/als etw.] (Lesart 2) sich entscheiden [für etw./jdn.] (E-Valbu, Lesart 1) gelten [für etw./jdn.] (Lesart 2)</p>	<p>Jdn./etw. halten [für jdn./etw.] (Lesart 1) heißen [für jdn.] (E-Valbu, Lesart 4 "für jemanden etwas zur Folge haben") (sich) interessieren [für etw.] (Lesart 1) kommunizieren [mit jdm.] mieten [für etw.] (Lesart 1) (z. B.: "Er mietet eine 2-Zimmer-Wohnung für 500€") passen [für XY] im Sinne von "geeignet sein für" (E-VALBU, Lesart 8) reden [mit jdm.], (reden, Lesart 6) sein [für etw./jdn.] (E-Valbu, Lesart 3) sorgen [für etw.] (Lesart 1) sprechen [mit jdm.] (sprechen, Lesart 4); ähnlich "kommunizieren [mit jdm.]" teilen [mit jdm.] (teilen, Lesart 4) (sich) treffen [mit jdm.] (sich treffen, Lesart 1) verbinden [mit etw.] (verbinden, Lesart 4) vergleichen [mit etw./jdm.] (vergleichen, Lesart 2) (sich) vorbereiten [für etw.] (Lesart 2) zahlen [für etw.] zusammenarbeiten [mit Kindern]</p>

<p>Objekt-PP Adjektiv</p>	<p>dankbar sein [für XY] (Duden Bedeutung 1; Ausdruck "dankbar sein für"; da Adjektivphrase, kein Attribut wie bei "Einfluss haben auf") geeignet sein [für etw.] empfänglich sein [für etw.] (Duden, Bedeutung 1 und 2) nötig sein [für etw.] (Duden, in Bedeutungserläuterung ist PP enthalten)</p>	<p>nützlich sein [für etw.] (Duden, in Bedeutungserläuterung ist PP enthalten) verantwortlich sein [für XY] (Duden Bedeutung 1) [für den Deutschen] wichtig sein offen [für XY] sein (Hinweis im Duden, Bedeutung 1a)</p>
<p>Adverbiale PP</p>	<p>[mit dem Auto] abholen arbeiten [mit etw./jdm.] arbeiten [für etw.] ("Er arbeitet für Firma x") (Lesart 2) aufgeschlossen sein [für XY] (PP nicht in Paraphrase im DUDEN enthalten, nur ein Beispiel) aussuchen [für XY] brauchen, es [für XY] (im Sinne von 'bedürfen': "Es braucht nicht lange, [...]", s. Duden Bedeutung 1c) brauchen [für etw.] (Lesart 2) X [mit dir] bringen [mit deiner Familie] fahren fahren [mit einem Auto] (s. E-Valbu, fahren II, kein PP-Komplement) (sich) freuen [für jdn./etw.] geben [für] (Adv. Bestimmung, weil E-Valbu "für-PP" nicht als Alternative zum Dativkomplement aufführt) gefährlich sein [für XY] (PP nicht in Paraphrase im DUDEN enthalten, nur ein Beispiel) genug sein [für X] ("genug" als Pronomen, s. Duden) glücklich sein [für XY] helfen [mit etw.] ideal [für mich] [für mich] interessant (PP nicht in Paraphrase im DUDEN enthalten, nur ein Beispiel) Eine Fahrkarte [für einen Wochenendausflug] kaufen (adv. PP s. E-Valbu, Lesart 1)</p>	<p>Glück [für die Geburt/die Prüfung] wünschen alles Gute/viel Glück [für XY] wünschen</p> <p>[für einen Termin] zur Verfügung stehen</p> <p>Das ist [...] attraktiv [für XY] einfach sein [für XY] gut [für uns] günstig [für mich] klein [für uns] neu [für mich] okay [für XY] richtig [für mich] schön [für mich] schwer [für mich] schwierig [für mich] sicher [für XY] sinnvoll [für XY]</p> <p>bekannt sein (für XY) (Keine PP im Duden)</p> <p>Anmerkung: Bei machen <i>für</i>-PP (<i>einfach sein für/günstig sein für/richtig für/schön für</i> etc.) ist der Status eines Adverbiale nicht abschließend geklärt. Jedoch finden sich zur Zeit der Annotationsphase in Referenzwerken (Duden-Online) keine</p>

	<p>[für die ganze Familie] kochen [mit großem Interesse] etw. lesen (Glückwünsche) senden [für XY] (für-PP kodiert Zielort, vgl. E-Valbu Lesart 3) etw. tun [für jdn.] X vorschlagen [für etw./jdn.] (im Sinne der Lesart 1 in E-Valbu)</p>	<p>eindeutigen Hinweise darauf, dass es sich um adjektivregierte <i>für</i>-PP handelt. Aus pragmatischen Gründen werden entsprechende <i>für</i>-PP konsistent als "Adverbiale PP" annotiert (und somit für die Auswertung der Fallstudie 3 nicht berücksichtigt).</p>
Attributive PP	<p>Antwort [für XY] Bewerbung [für XY] ("Bewerbung für ein Praktikum") Dokumente [für Visum ...] Einladung [für das Gespräch] Interesse [für etw.] haben/wecken Kontakt [mit Leuten und Kindern] lieben (Beispiel aus B1-Niveau) Problem [mit XY] (im Sinne einer bestimmten Schwierigkeit) X [mit Schwerpunkt Y]</p>	<p>Termin [mit meinem Chef] Voraussetzung [für die Stelle] erfüllen/mitbringen Zeit [für ein Gespräch] (OA) Zeit [für XY] haben (ähnlich zu "Einfluss bekommen [auf die Gesellschaft]", Duden, S. 791) Zusammenarbeiten [mit anderen] etw. [mit Mozart] haben (i.S.v. ein Produkt mit Mozart-Bezug)</p>
Rest	<p>Ist alles ok [mit ihr]? Mit wie vielen kg ist es geboren? (Herzlichen) Glückwunsch [für XY] (-> laut DVDs keine typische Verbindung, dennoch formelhafter Charakter der PP; unsicher)</p>	
Pseudotreffer	<p>Im Vergleich mit (präpositionsartige Fügung) Mit Annahme von.. (präpositionsartige Fügung) Mit Bezug auf .. (präpositionsartige Fügung) Mit Hilfe (von) (mithilfe als sekundäre P) mit teilnehmen an (eine Art Verbpartikel) Ein/Das Für und Wider (Präposition ist substantiviert, s. auch Duden Bedeutung 1b)</p>	
Formelhafte Sequenz	<p>da sein [für jdn./etw.] (in E-VALBU Hinweis auf Idiom. Wendung "da sein" i.S.v. "behilflich sein") [Mit vielem Dank] im Voraus (o.ä.)</p>	<p>(Vielen) Dank [für XY] Danke [für XY] (ohne Subjekt) Danke sehr [für XY] Danke schön [für XY]</p>

	<p>[Mit dem Gruß] [Mit freundlichem GruÙe] [Mit freundlichen GrüÙen] [Mit Gruß] [Mit schönem Gruß]</p> <p>Viele GrüÙe [für XY] (ohne Verb im Satz, am Ende des Textes als Formel, ähnlich zu "Mit freundlichem Gruß")</p> <p>mit sich bringen ("bringen", Duden, Bedeutung 4, unter "Wendung") es mit etw./jdn. zu tun haben ("tun", Duden, Bedeutung 7) verbleiben [mit Hoffnung/GrüÙen/o.ä.]</p> <p>eine Rolle [für jdn./etw.] spielen (s. Duden "Rolle", unter "Wendungen/Redensarten/Sprichwörter")</p>	<p>Dankeschön [für XY]</p> <p>(Viel) Glück [für XY] (ähnlich zu (Vielen Dank für XY, s.o.)</p> <p>Was [für ein] / Was [für] u. ä. / Was gibt es [für XY] (Wendung bei Duden-Online, unter "für")</p> <p>Wie geht es [mit X]? Wie ist es [mit X]? Wie wäre es [mit X]?</p> <p>"Zwillingsformel": X für X ("Tag für Tag") (nach Helbig/Buscha 2001b)</p>
--	---	--

Dokument Beispiele zur Annotation von PP-Kontexten des Fehlertyps Hinzufügung

Beispiel aus MERLIN	Hinweise zur Annotation
<p>Tok: da meinen Kindertraum ist , mit fremde Leute kennen zu lernen und fremde Länder zu besuchen TH1: da es mein Kindertraum ist , fremde Leute kennen zu lernen und fremde Länder zu besuchen</p>	<p>Annotation: Objekt-PP Hier wurde PP zu NP geändert. PP in tok = mit fremde Leute NP auf TH1= fremde Leute - „fremde Leute“ ist Objekt zum Verb „kennnenlernen“</p>
<p>Tok: entweder Sie hier allein oder mit Ihrer Familie mitbringen wollen , TH1: ob Sie entweder hier allein sind oder Ihre Familie mitbringen wollen , dann sollen Sie ein</p>	<p>Annotation: Objekt-PP Hier wurde PP zu NP geändert. PP in tok = mit Ihrer Familie mitbringen NP auf tok= Ihre Familie mitbringen - „Ihre Familie“ ist Objekt zum Verb „mitbringen“</p>
<p>Tok = Ich weiss , dass ich mit den Kindern in einen Deutschefamilie pflegen muss TH1: Ich weiss , dass ich die Kinder in einer deutschen Familie pflegen muss</p>	<p>Annotation: Objekt-PP Hier wurde PP zu NP geändert. PP auf tok = mit den Kindern in einen Deutschefamilie pflegen muss TH1 = die Kinder in einer deutschen Familie pflegen muss - „die Kinder“ ist Objekt zum Verb „pflegen“</p>

